Kinetic Barriers to Folding in the *Tetrahymena* Group I Ribozyme

by

Martha S. Rook

B.S., Chemistry
Texas A&M University, 1992

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
IN BIOCHEMISTRY

at the

Massachusetts Institute of Technology

June, 1998

Signature of Author..... .....
5/13/98

Certified by.....
James R. Williamson
Professor of Chemistry
Thesis Supervisor

Accepted by.....
Dietmar Seyferth
Chair, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor Lawrence J. Stern.....
Chair

Professor James R. Williamson.....
Thesis Supervisor

Professor John M.Essigmann.....

Kinetic Barriers to Folding in the *Tetrahymena* Group I Ribozyme

by

Martha S. Rook

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biochemistry

Abstract

The process by which RNA molecules fold into complex three dimensional structures is a fundamental question in biochemistry. The kinetic folding of the *Tetrahymena* ribozyme has been investigated and a model for the kinetic folding pathway has been proposed where the two main structural domains, P4-P6 and P3-P7, form in a hierarchical manner with P4-P6 forming first and P3-P7 folding on the minute timescale. Recent studies identified a set of mutations that accelerate P3-P7 formation, and all of these mutations appear to disrupt a kinetic trap stabilized by native contacts in P4-P6. To better understand the microscopic details of the slow folding step, a previously developed kinetic oligonucleotide hybridization assay was used to characterize the folding of several fast folding mutants. A comparison of the temperature dependence of the wild type and mutant ribozyme P3-P7 folding rates demonstrates that a majority of the mutations act by decreasing the activation enthalpy required to reach the transition state. The denaturant urea accelerates the rate of folding of the wild type, while for the mutants urea merely shifts the distribution between two folding populations. The mutations do not accelerate formation of the catalytically active structure, suggesting that there is an additional slow step on the folding pathway. Comparison of the Mg^{2+} dependence of the two slow steps provides additional evidence that they are distinct and suggests that the final slow step represents escape from a magnesium stabilized off-pathway intermediate. The work presented in this thesis supports the emerging view of RNA folding as an escape from a series of kinetic traps and highlights the dual role of Mg^{2+} in stabilizing both the final native structure and off-pathway folding intermediates that slow folding. The analysis of the fast folding mutants demonstrates that small structural changes or changes in solvent can accelerate folding, but these changes lead to complex folding behavior, and do not give rise to rapid two-state folding transitions. These results support the recent view of folding as an ensemble of molecules traversing a rugged energy landscape to reach the lowest energy state.

Thesis Supervisor: James R. Williamson

Title: Professor of Chemistry

Acknowledgments

The course of my graduate work has taken place over six years, and in those six years a large number of people helped me greatly and contributed to all the things you must learn to get a Ph.D.. First and foremost is my advisor Jamie Williamson. Jamie has been a wonderful advisor to work for, he has a boundless excitement about science and RNA that ends up being transmitted to his students. He has always been able to walk the line between giving his students independence and being there to give advice when necessary. Jamie has also created a very good lab atmosphere full of both scientific collaboration and social friendship.

I would also like to thank my two collaborators on the RNA folding project, Pat Zarrinkar and Dan Treiber. Pat was the graduate student who first trained me to work with RNA, and it was his guidance and meticulous attention to detail that taught me how to avoid RNase contamination. Pat was also the originator of the RNA folding project and without him very little work in this thesis would exist. Dan Treiber joined the RNA folding project as a post-doc shortly before Pat graduated, and shared many a conversation about folding energy landscapes, the theoretical basis of the effects of urea (no one really knows!), and what constitutes an on-pathway folding intermediate. I would especially like to thank Dan for reading my manuscripts and giving helpful suggestions while overlooking many of my grammatical and spelling errors. I'd like to thank all of the members of the Williamson lab for making up a friendly and always entertaining lab atmosphere. The lab always had something fun and crazy going on from lab seminars on knot tying led by Jeff Orr to the Funke factoid of the day. There were always labmates ready to make a Tosci's run or to head to the Muddy to let of steam.

Graduate school wouldn't have been the same without the lunch bunch Evan Powers, Ben Turk, Jeff Ekert, Dave Wang, Bill Kobertz, and Mat Bogio who gave me a reason to get out of the lab for at least an hour every day. They also were largely responsible for my involvement in

chemistry summer league volleyball (such as it is), even though they soon were far beyond me in skill. Thanks to Evan Powers for getting me addicted to the X-files and South Park and for hosting many a party to encourage viewing, to Bill Kobertz for introducing me to Jackie Chan, and to Ben Turk for introducing me to indie rock.

I'd like to thank my parents, Lois and Harry Rook, who have always supported me and taught me that I could do anything I put my mind to. It is through a combination of their good genes and the nurturing environment they provided during my childhood that have allowed me to pursue my goals and always follow my dreams.

Finally I'd like to thank John Bearley, who six years ago decided to move from Texas to Massachusetts with me. He has been a constant source of love and support and made the process of being a graduate student much more enjoyable. I've been lucky to have him here to remind me that other things in life than grad school are important, to drag me away for trips to Colorado and other fun locations, and in general to be my best friend and companion in life.

Table of Contents

1. The RNA Folding Problem

The many cellular roles of RNA.

Until the discovery of RNA catalyzed reactions (Cech *et al.*, 1981; Guerrier-Takada & Altman, 1984; Guerrier-Takada *et al.*, 1983; Kruger *et al.*, 1982), RNA was seen as a passive information carrier. In the decade since this discovery, it has been realized that RNA plays an active role in a number of biological processes from RNA processing (Jacquier, 1997) to perhaps even catalyzing peptide bond formation during translation (Noller, 1991) (Figure 1). Indeed, RNA is active in almost all aspects of gene expression, and *in vitro* selection experiments have suggested that RNA may be able to perform a variety of additional functions as well as those seen in the cell.

Genomic Maintenance

RNA is involved in maintaining the linear ends of chromosomes through the ribonucleoprotein (RNP) enzyme telomerase. Telomeres are the ends of eukaryotic chromosomes that stabilize the chromosome (Blackburn, 1991), and consist of tandemly repeated sequences of DNA rich in guanosines. Telomeres are progressively shortened by cell division, and telomerase maintains the length of telomeres through an RNA templated reverse transcription process that synthesizes the telomeres. The telomerase RNA acts as the template for telomere synthesis, and the protein component acts as a reverse transcriptase.

Gene Expression

Transcription is the first step on the path leading from an encoding sequence to a functional protein, and proteins play the dominant role in this process. However, there are instances where RNA plays a role in enhancing processivity and anti-termination. One example is the HIV TAR element which binds to the Tat protein and increases transcription processivity for the viral RNA (Frankel, 1992).

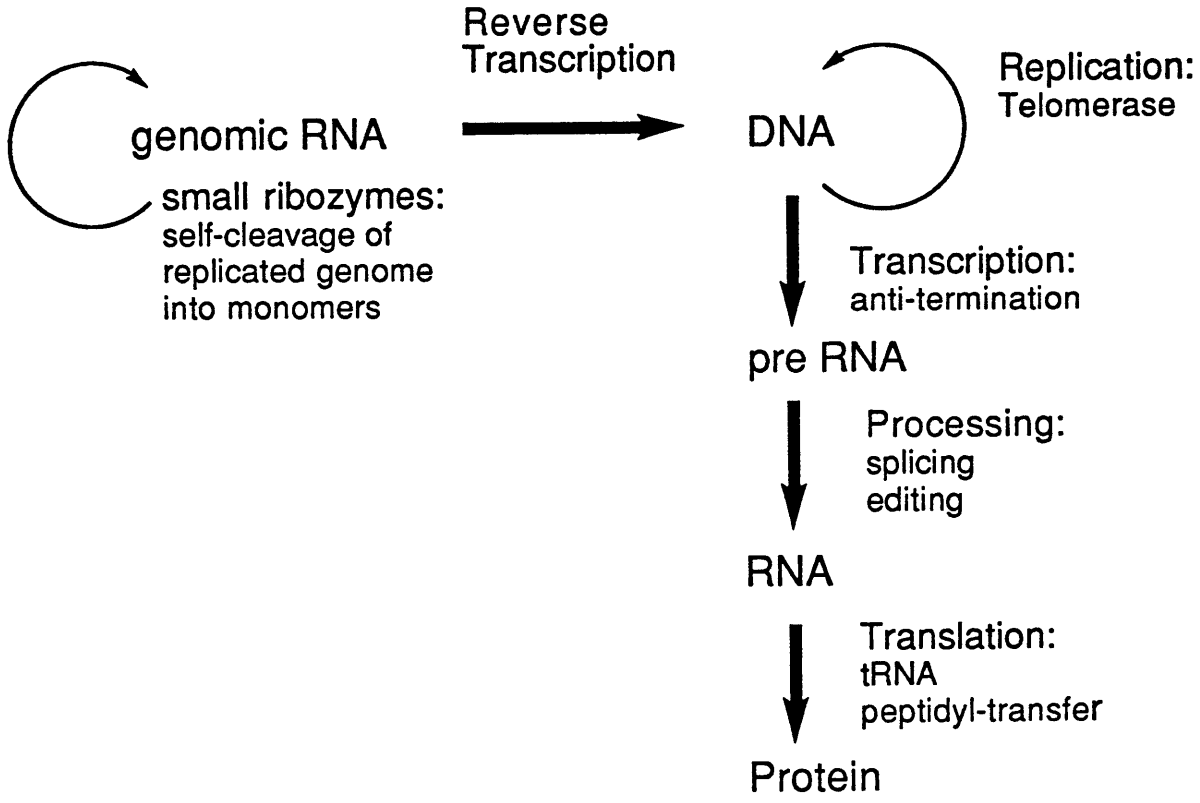


Figure 1: The central role of RNA.

RNA processing is generally required before translation can occur, and RNA often plays a central role in its own modification. It is most predominantly active in the removal of intervening sequences, the reaction ribozymes were first discovered to catalyze. Intervening sequences, or introns, are sequences that interrupt coding sequences of genes, and are removed by the process called splicing. RNA introns that catalyze their own excision from primary transcripts can be grouped into four major categories based on structural features such as sequence and tertiary fold and by their catalytic mechanisms. These categories are the group I introns, which are widespread and found in mRNA, rRNA and tRNA genes in the mitochondrial, chloroplast, and nuclear genomes of eukaryotes, bacteriophages, and eubacteria, but not in vertebrates; nuclear pre-tRNA introns; the group II introns which are mainly found in eukaryotic organelles; and nuclear pre-mRNA introns.

Pre-mRNA introns in eukaryotes are removed by the spliceosome. The spliceosome is a large, dynamic complex consisting of five small RNAs, U1, U2, U4, U5, and U6, which form RNP complexes with over 50 proteins. The spliceosomal RNAs are believed to play the central role in both spliceosome assembly and catalysis (Guthrie, 1991; Steitz, 1992). The group II introns and the nuclear spliceosome share very similar RNA structure and splicing mechanism, which has lead to the proposal that the two are evolutionarily related (Madhani & Guthrie, 1994; Moore, 1993; Yu *et al.*, 1995).

While some introns catalyze their own excision, they are not technically true enzymes as they are not capable of multiple turnover. While these introns can often be modified to make them capable of multiple turnover, Ribonuclease P (RNase P), a ribonucleoprotein complex that catalyzes the removal of the 5' terminal leader sequence of tRNAs, is a true ribozyme (Altman *et al.*, 1995). While RNase P does contain protein as well as RNA, it has been shown that the RNA is the catalytic molecule and can catalyze the cleavage of pre-tRNA in the absence of the protein component *in vitro* (Guerrier-Takada et al., 1983).

In addition to intron sequences that catalyze their own excision, and RNase P, there are a number of small ribozymes generally found in small plant viruses. These include the hammerhead ribozyme, hairpin ribozyme, and hepatitis delta ribozyme (Symons, 1992). These viruses replicate via a rolling circle mechanism that generates long RNAs that are cleaved by the ribozymes into monomers.

While many of these ribozymes cleave RNA via different mechanisms, and some require protein cofactors while others do not, there is a common theme among them. All of the ribozymes must fold into fairly complex structures and require divalent metal ions both to stabilize structure and to aid in the cleavage reaction, and in all cases the RNA molecule is primarily responsible for catalysis.

In addition to splicing, RNA editing can occur to post-transcriptionally modify RNA. The editing usually consists of the insertion of additional bases (Simpson & Thiemann, 1997). The catalytic mechanism has not yet been determined, but it may be catalyzed by protein enzymes but it has been suggested that small RNAs called guide RNAs may also catalyze the reaction using a mechanism similar to that of group II introns. In support of this mechanism, gRNA-mRNA chimeric molecules have been isolated (Blum *et al.*, 1991).

RNA is also centrally involved in the translation process. The ribosome is a large ribonucleoprotein complex with three RNA domains and a multitude of ribosomal proteins. RNA is involved on many levels in this process. Transfer RNA translates the three base codons to their corresponding amino acids, and the structure and function of tRNA has been extensively investigated (Giegé *et al.*, 1993; Nissen *et al.*, 1995; Rould *et al.*, 1991). The ribosomal RNA contacts the anti-codon region of tRNA (Moazad & Noller, 1990; Moazed *et al.*, 1986), and the tRNA acceptor stem (Barta *et al.*, 1984; Moazed & Noller, 1989) forming the active site for peptidyl transfer. In addition rRNA has been shown to bind both elongation factors EF-Tu and EF-G (Moazad *et al.*, 1988). Ribosomal 16S RNA is known to contact the Shine-Delgarno sequence of mRNA to align the mRNA on the ribosome during initiation (Jacob *et al.*, 1987; Shine

& Dalgarno, 1974). Genetic experiments (Dabbs, 1991) and *in vitro* experiments (Franceschi & Nierhause, 1990; Noller *et al.*, 1992; Schulze & Nierhaus, 1982) showing that most ribosomal proteins are not required for peptidyl transfer have led to the proposal that the RNA is in fact responsible for catalysis.

In vitro RNA evolution

There has long been speculation about the existence of an RNA world, a prebiotic world where RNA was the sole genetic material and catalyzed its own replication (Crick, 1968; Orgel, 1968; Woese, 1967). While evidence of this time period no longer exists, evolution in the test tube is one method used to gain evidence that RNA based information transfer is feasible. Such experiments called *in vitro* selection or evolution can also be used for a rapidly expanding number of purposes. Selection techniques have been used to better define protein-RNA interactions such as tRNA binding to tRNA-synthetase (Peterson *et al.*, 1994) and the spliceosomal protein U1A to U1 small nuclear RNA (Laird-Offringa & Belasco, 1995), as well as to find RNAs that bind tightly and specifically to small molecules such as ATP (Sassanfar & Szostak, 1993), and arginine (Famulok, 1994), and even to discriminate between such similar molecules as theophylline and caffeine (Jenison *et al.*, 1994). Selection can also be used to construct artificial phylogenies to confirm secondary structure predictions when few sequences are known (Berzal-Herranz *et al.*, 1992). Selection experiments have also been used to better understand the structural requirements of existing ribozymes (Couture *et al.*, 1990; Doudna & Szostak, 1989; Green & Szostak, 1994), and to evolve novel ribozymes with new functions (Beaudry & Joyce, 1992; Ekland & Bartel, 1996; Wilson & Szostak, 1995).

In vitro selection has proven to be a powerful tool in exploring RNA structure and function. The new RNAs evolved give insight into common RNA folding motifs and help in the understanding of the mechanisms by which RNA can catalyze a variety of reactions. Selection techniques have also been used to design RNAs with specific characteristics for therapeutic purposes. While the selected RNAs do not necessarily have counterparts in the cell, they do shed

light on RNA structure and function. New catalytic RNAs have shown the ability of RNA to catalyze reactions beyond the simple cleavage and formation of phosphodiester bonds and such ribozymes may one day be found in natural systems.

In the decade since catalytic RNA molecules were discovered, much work has been done to show that RNA is not the passive molecule it was once thought to be. It has been shown to play a major role in almost all aspects of information management in the cell, and has been found to be able to catalyze a number of additional reactions through targeted selection. There are still a number of questions to be answered about RNAs function in large RNP complexes such as the ribosome and spliceosome and may be other as yet unknown RNA functions in the cell.

RNA Structure and Folding

In order for RNA to carry out the many functions described above it must fold into complex three dimensional structures. Understanding these structures and the processes by which they form may provide insights into how they are able to carry out their function. Therefore there is an RNA folding problem analogous to the much studied protein folding problem. What are the structures that RNA can form? Are there common folding motifs, and how are these motifs stabilized? What roles do cofactors such as divalent metal ions and proteins play in structural stabilization? In addition, the process by which these structures are formed must be understood. This requires the understanding of the pathways by which these molecules fold, analysis of any intermediates on the pathway, and the role that metal ions play.

Small RNA structures.

Large RNA structures have been difficult to study by either NMR or X-ray crystallography. This has lead to the study of small RNA motifs as model systems for understanding general themes in RNA folding. The structures of small RNA helices (Baeyens *et al.*, 1996; Limmer *et al.*, 1996; McDowell & Turner, 1996; Wu & Turner, 1996) and hairpin loops (Fountain *et al.*, 1996; Mirmira & Tinoco, 1996) have shown the importance of stacking and non-canonical base-pairing to structure stabilization. The folding of RNA duplexes and small helices

has also been studied and they have been found to fold on the microsecond to millisecond timescale (Gralla & Crothers, 1973; Pörschke & Eigen, 1971). The structures of small folding motifs such as pseudoknots (Puglisi *et al.*, 1990; Shen & Tinoco, 1995) and hairpin loop-loop interactions (Marino *et al.*, 1995) have been solved, and both structures require significant bending necessary for their biological function. The structures of RNA aptamers, RNA molecules selected *in vitro* to bind specific small molecules, are interesting because they can be studied in the presence and absence of the molecules they bind. This gives information about how conformational changes are important to the binding process. A common fold in aptamers consists of a hairpin containing an internal loop that contains the binding site (Dieckmann *et al.*, 1996; Yang *et al.*, 1996). The internal loop is unstructured in the absence of the small molecule but becomes structured upon binding. The structures of a number of RNA-protein complexes have been solved using the minimal RNA sequence required for binding (Battiste *et al.*, 1996; Puglisi *et al.*, 1995). These structures have given insights into both RNA structures and protein induced conformational changes. Similar to aptamer binding of small molecules, the RNA-protein binding site usually involves an irregular RNA helix with a widened major groove.

The structures of these small RNA motifs give molecular details of the secondary structures possible in folded RNA. To understand higher order folding, the structures of more complex RNAs must be solved. The hammerhead ribozyme is a small self-cleaving ribozyme from a small RNA plant virus. The ribozyme is approximately fifty nucleotides in size, and consists of three helices flanking a phylogenetically conserved core. The crystal structure of the hammerhead ribozyme has been solved with an all DNA (Pley *et al.*, 1994b) and all RNA (Scott *et al.*, 1996) substrate strand as well as complexed with an RNA strand containing a 2'-O-methyl modified RNA (Scott *et al.*, 1995) at the cleavage site that acts as an inhibitor (Figure 2A). The fold of the hammerhead ribozyme is shaped like a wishbone with stems I and II forming the arms and stem III and the core forming the base. A tight turn, called the U-turn due to the conserved CUGA sequence involved, is nearly identical to a uridine turn found in the structure of phenylalanine

tRNA (Kim *et al.*, 1974). The cleaved phosphodiester bond is not positioned correctly for cleavage in the structure and it has been proposed that a conformational change must occur to achieve the transition state. These results highlight conserved folding motifs in RNA structures, and the conformational flexibility of RNA structures that appears necessary for RNA function.

The structure of tRNA was one of the first RNA structures to be solved. Transfer RNA is a ubiquitous molecule and is found in all cell types, and perhaps the ease of isolating this molecule led to it being one of the earliest and most extensively studied RNAs. A secondary structure has been determined, and an X-ray crystal structure has been solved (Kim et al., 1974; Robertus *et al.*, 1974). tRNAs have a common cloverleaf secondary structure and tertiary fold (Figure 2B) composed of four short helices the D-stem, T-stem, acceptor stem and anticodon stem. The native structure can form in the presence of sodium ions but is greatly stabilized by the addition of magnesium. In solution tRNA has a great deal of conformational flexibility and it has been proposed that this could help mediate it's binding to aminoacyltransferases and the ribosome (Crothers & Cole, 1978).

The folding and magnesium binding of tRNA has been studied extensively using thermal unfolding (Cole *et al.*, 1972), temperature jump relaxation kinetics (Cole & Crothers, 1972), NMR (Crothers *et al.*, 1974), fluorescence (Lynch & Schimmel, 1974), and calorimetry (Privalov *et al.*, 1975). Based on these studies both an equilibrium and a kinetic folding pathway has been proposed. Secondary structure forms first and in the presence of only sodium ions occurs in a stepwise fashion, when magnesium is present structure formation is more cooperative. The rate limiting step when folding is initiated by the addition of magnesium at intermediate sodium concentrations (170 mM Na^+) is formation of tertiary structure, which occurs on the millisecond timescale. When folding is initiated from low salt conditions (12 mM Na^+), the rate of folding is much slower, on the second timescale. The high activation enthalpy for the folding led to the proposal that under low salt conditions misfolded secondary structure forms and must be disrupted before productive folding could occur.

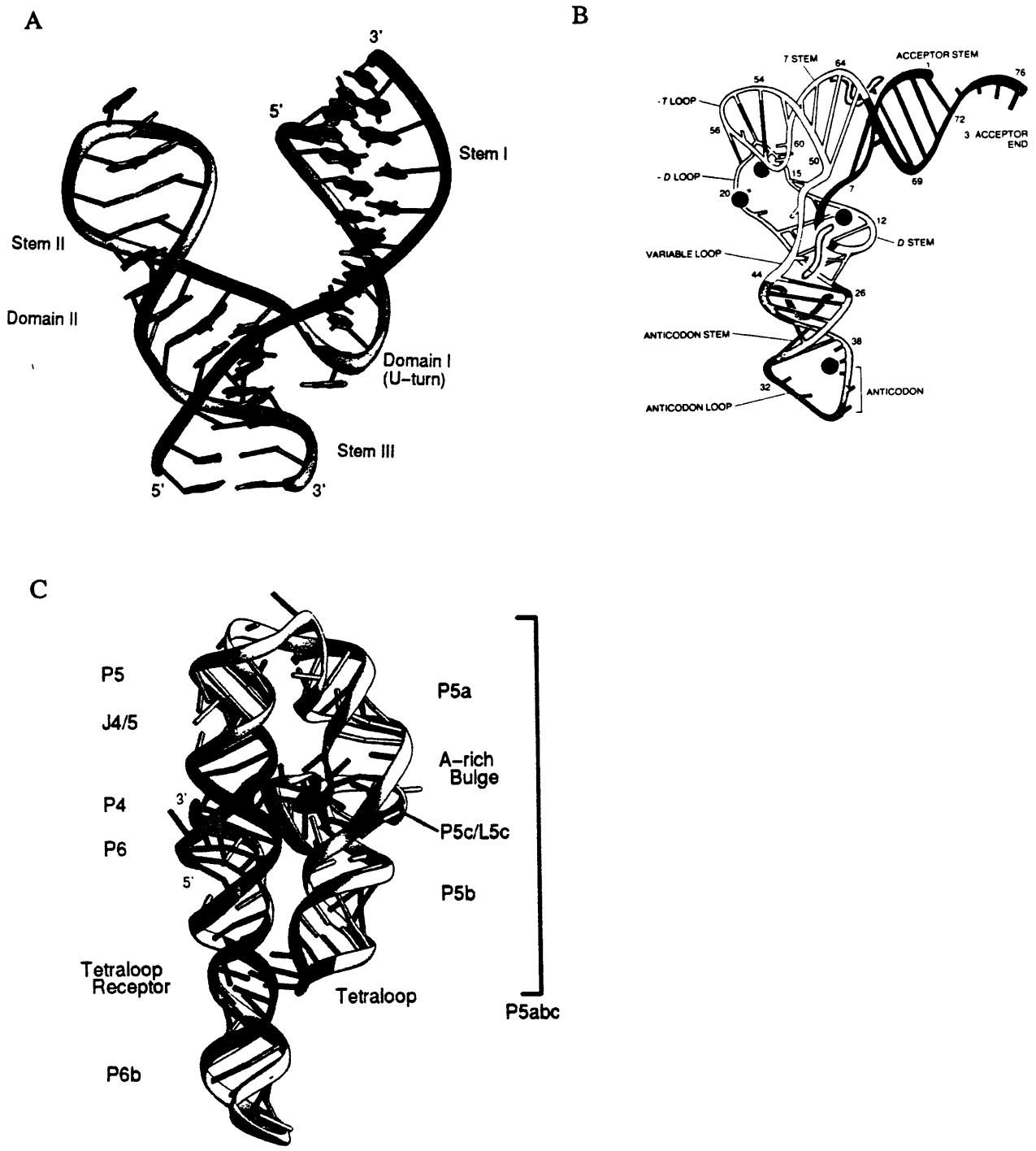


Figure 2: Three dimensional structures of three RNAs. A. The hammerhead ribozyme from (Scott et al., 1995) B. Model of tRNA tertiary structure from (Söll, 1993) C. Structure of the *Tetrahymena* ribozyme P4-P6 domain from (Cate *et al.*, 1996a).

Large, highly structured RNAs

The structures and folding pathways of large, highly structured RNAs is much more complicated than those discussed above. No three dimensional structures have been determined by NMR or crystallography for large RNAs (> 200 nucleotides), although a structure of a domain of the *Tetrahymena* ribozyme has been solved (Figure 2C) (Cate et al., 1996a). Models of the three dimensional structures of large ribozymes such as RNase P (Harris *et al.*, 1994; Westhof & Altman, 1994), the *Tetrahymena* ribozyme (Harris et al., 1994; Lehnert *et al.*, 1996; Michel & Westhof, 1990; Westhof & Altman, 1994), and 16S ribosomal RNA (Stern *et al.*, 1988) have been proposed based on both biochemical techniques and phylogenetic analysis, these models show that large RNAs tend to have discrete folding domains. Analysis of equilibrium magnesium binding pathways supports this view with equilibrium folding occurring in a hierarchical manner (Celander & Cech, 1991; Pan, 1995), secondary structures close in sequence forming first followed by long range interactions and formation of tertiary structure. Mapping of the tertiary structures using Fe(II)-EDTA generated hydroxyl radicals measures the solvent accessibility of the RNA backbone and has shown that in the presence of divalent metal ions such as magnesium, large RNAs have a closely packed structure with an inside and an outside (Celander & Cech, 1991; Latham & Cech, 1989; Pan, 1995).

The folding of a few of these large RNAs has been analyzed. Both the equilibrium and kinetic folding of RNase P have been studied. RNase P is a ubiquitous ribozyme that catalyzes the endonucleolytic removal of the 5' leader sequences from pre-tRNA generating mature 5' ends. It is a ribonucleoprotein complex containing an approximately 400 nucleotide RNA domain and an approximately 120 amino acid protein domain. Both domains are needed for *in vivo* activity, but *in vitro* at elevated magnesium concentrations the RNA domain alone can catalyze the cleavage reaction (Guerrier-Takada & Altman, 1984; Reich *et al.*, 1988). The protein component is thought to facilitate helix packing, and accelerate product release (Reich et al., 1988). While RNA activity

requires high magnesium concentrations (20 mM) folding is complete at approximately 2-3 mM Mg^{2+} (Pan, 1995), therefore the protein component is not required for productive RNA folding.

The kinetic folding of RNase P RNA has been studied using both kinetic oligonucleotide hybridization (Zarrinkar *et al.*, 1996) and circular dichroism (Pan & Sosnick, 1997). Oligonucleotide hybridization experiments targeted specific helices with complementary DNA oligonucleotides and measured the accessibility to binding at various times after folding was initiated by the addition of magnesium. In the absence of magnesium a number of short range helices are already formed but at least one helix (the P7 helix) requires magnesium to form and folds with a slow rate (0.5 min^{-1}). Experiments monitoring equilibrium folding using circular dicroism agreed with these results and found that at low magnesium concentrations (0.4 mM Mg^{2+}) helix formation occurs and higher magnesium concentrations (2 mM) are required for tertiary structure formation. The kinetic folding was also measured and correlated with the changes seen in the CD spectrum during equilibrium folding. The first folding phase occurs on the millisecond timescale, and is believed to correspond to secondary structure formation. There is a second slow step that occurs on the minute timescale and corresponds to formation of tertiary structure, this could also correspond to the formation of P7 measured by oligonucleotide hybridization. The slow step was found to have a high activation enthalpy (56 kcal/mol) and was accelerated by the addition of the denaturant urea. Therefore, it was proposed to be the result of a kinetic trap.

The ability of misfolded structures to slow RNA folding is highlighted in the study of the folding of pre-RNA from the *Tetrahymena* group I intron. The pre-RNA 5'-exon sequence can misfold after transcription, and form a helix which inhibits formation of the 5' splice site and slows folding to the catalytically active structure (Emerick & Woodson, 1993; Emerick & Woodson, 1994). Native polyacrylamide gel experiments monitoring the conformations of RNA during folding have showed that a large population of misfolded RNA exists after transcription and is slowly converted to the active structure. Iterative annealing can increase the population of correctly

folded RNA, and addition of urea can accelerate the folding process. Therefore, it has been proposed that the folding pathway of the pre-RNA involves multiple parallel pathways resulting from the rapid collapse of the RNA into misfolded states that must be disrupted for productive folding (Pan *et al.*, 1997).

Protein chaperones in RNA folding.

RNA misfolding appears to be a common problem in RNA folding and is due to the thermodynamic stability of alternative secondary structures and tertiary interactions. Nonspecific RNA binding proteins have been shown to facilitate splicing *in vitro* (Coetzee *et al.*, 1994; Herschlag *et al.*, 1994; Zhang *et al.*, 1995a), these proteins most likely act by disrupting misfolded structures and allowing them another chance to fold to the correct structure (Herschlag, 1995). Proteins can also specifically bind to RNA structures to promote folding to the biologically active structure. The binding of the human U1A protein to model helices containing an internal loop causes a sharp bend in the helix (Allain *et al.*, 1996). The binding of the S15 ribosomal protein to 16S rRNA properly aligns a three helix junction, high magnesium concentrations can substitute for the protein and allow alignment of the junction (Batey & Williamson, 1996). The structures and folding pathways of a number of group I introns have been studied, the *Tetrahymena* ribozyme has been the most extensively studied and is reviewed in the next section. Group I introns have a common folded structure composed of two domains P4-P6 and P3-P7 which together form the catalytic core of the ribozyme (Cech, 1993). Many group I introns have a variety of peripheral domains that stabilize the core and assist in folding and some introns have protein cofactors instead of peripheral RNA domains that appear to play the same role (Lambowitz & Pearlman, 1990). The splicing of the *Tetrahymena* ribozyme is more rapid *in vivo* that *in vitro* (Zhang *et al.*, 1995b), leading to the suggestion that unidentified protein cofactors assist the ribozyme in cellular splicing.

The folding of two group I introns that require protein cofactors has been studied in detail, the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase (CYT-18 protein) facilitates splicing of the ND1 intron (Guo & Lambowitz, 1992; Mohr *et al.*, 1994; Mohr *et al.*, 1992; Saldanha *et al.*,

1995) and CBP2 protein facilitates splicing of a yeast mitochondrial group I intron (bI5) (Weeks & Cech, 1995; Weeks & Cech, 1996). Although both proteins stabilize the catalytic core of the ribozyme they do so in different ways. In the absence of the CYT-18 protein only secondary structure is formed in the ND1 intron, the protein binds to the P4-P6 domain stabilizing the structure and scaffolding P3-P7 formation (Caprara *et al.*, 1996), much like the P5abc domain in the *Tetrahaymena* ribozyme (Cate *et al.*, 1997; Mohr et al., 1994). In contrast, in the bI5 intron folding of the core occurs transiently and is trapped in the folded state by the binding of the CBP2 protein. The CBP2 protein can only bind the properly folded intron and its main function is to stabilize this structure (Weeks & Cech, 1996).

Common themes in RNA folding.

The studies of RNA folding detailed above highlight some common themes in RNA folding (Figure 3). RNA secondary structure is stabilized by both stacking and base-pairing, and RNA often adopts non-Watson-Crick base-pairing making RNA helices more irregular than DNA helices. Tetraloops and internal loops are important in the binding of small molecules and proteins by RNAs. RNAs can also form complex tertiary structures with compact folded domains much like proteins can. While protein folding appears to be driven by the collapse of hydrophobic residues, it appears that in RNA magnesium may facilitate packing of the phosphate backbone and drive the folding process. Proteins can both substitute for magnesium to facilitate this packing, and also act as RNA chaperones to minimize misfolding and stabilize correct structures. Folding appears to be hierarchical with more stable secondary structures forming first, followed by formation of tertiary structure. However, misfolded kinetic traps appear to be an important factor in slowing folding.

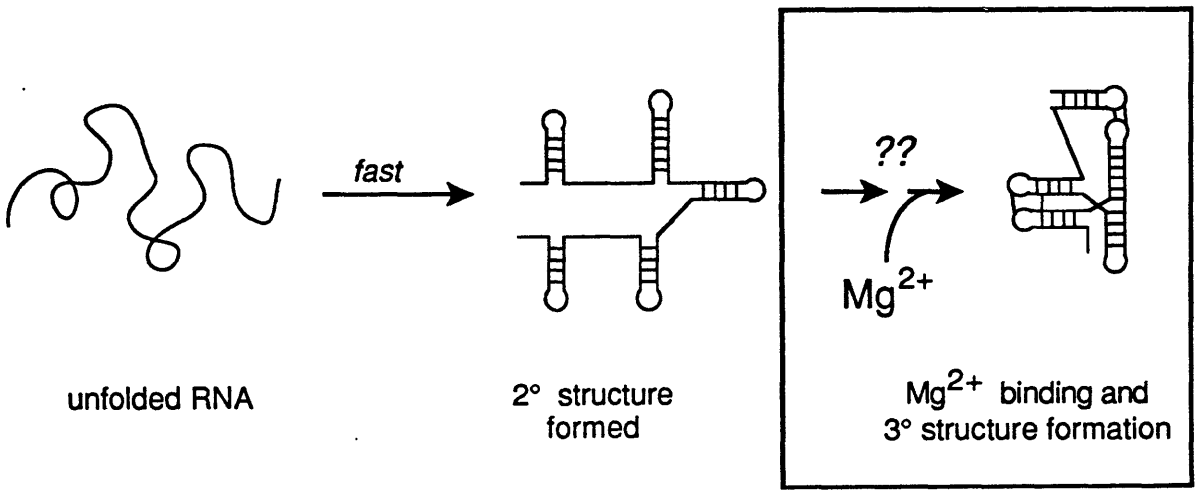


Figure 3: A minimal pathway for RNA folding.

The *Tetrahymena* ribozyme as a model system for studying RNA folding.

The *Tetrahymena* group I ribozyme was the first catalytic RNA to be discovered (Cech et al., 1981; Zaug & Cech, 1982), and as such it is one of the best understood catalytic RNAs. The self-splicing reaction catalyzed by the *Tetrahymena* ribozyme has been well studied, and a mechanism proposed. A model for the three dimensional structure has been proposed based on extensive biochemical analysis and phylogenetic comparison (Lehnert et al., 1996; Michel & Westhof, 1990). The detailed information about the structure and function of this catalytic RNA makes it an ideal system for studying RNA folding.

Function

The *Tetrahymena* group I intron was first discovered through studies of intervening sequences in pre-rRNA. A catalytic activity that removed the intron from rRNA exons could not be purified from the RNA fraction, this catalytic activity turned out to be the RNA itself and the first ribozyme was discovered (Cech, 1990). Since its discovery, much work has been done to determine the mechanism of the self-splicing reaction. The self-splicing proceeds by two consecutive transesterification reactions. The first transesterification occurs when a guanosine bound to the ribozyme attacks the phosphorus atom of the 5' splice site and forms a 3',5'-phosphodiester bond to the first nucleotide of the intron. This leaves a free 3' hydroxyl at the 5'-splice site and after a conformational rearrangement this hydroxyl attacks the 3' splice-site leading to ligated exons and an excised intron which can then undergo further cyclization reactions (Cech, 1993).

A number of cofactors are necessary for self-splicing to occur. Divalent metal ions are necessary for structure formation and activity, and the metal requirement is usually filled by Mg^{2+}. The other divalent metals Mn^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} can relieve the Mg^{2+} requirement for structure formation, but only Mn^{2+} can substitute for Mg^{2+} in terms of the activity requirement.

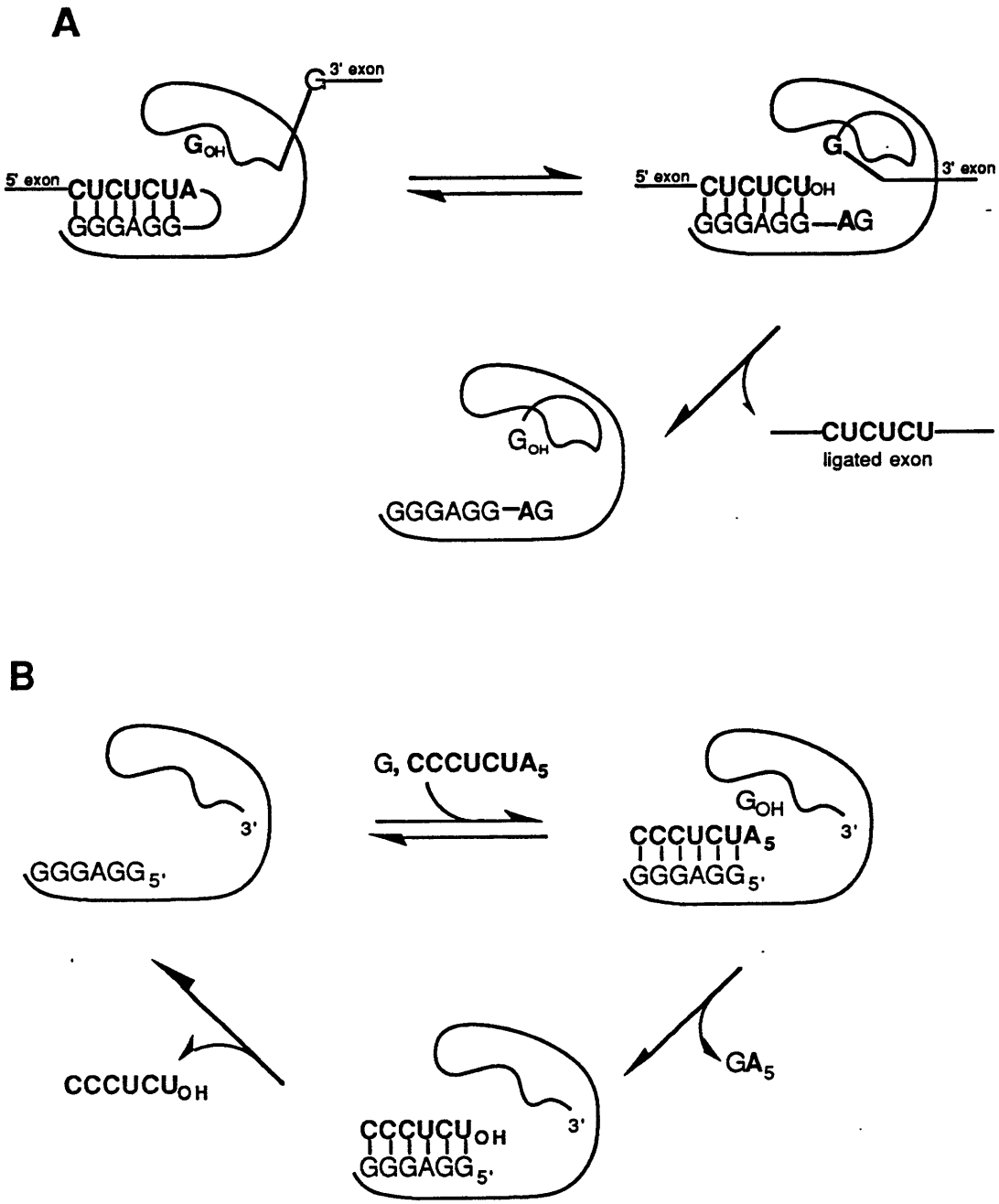


Figure 4: *Tetrahymena* riboyzme function. Adapted from Herschlag and Cech (1990a).

A. self-splicing. B. endonuclease cleavage.

For optimal activity, guanosine, or one of its 5'-phosphorylated forms (GMP, GDP, or GTP) is required. However, in the absence of guanosine a less efficient hydrolysis reaction can occur.

The *Tetrahymena* ribozyme has a short duplex (4-7 nts) called the P1 helix (Been & Cech, 1986; Waring *et al.*, 1986; Zaug *et al.*, 1986) that specifies the 5' splice site. The 3' sequence of this helix is also called the internal guide sequence (IGS) as it properly aligns the P1 helix in the substrate binding pocket and sets up cleavage at the appropriate phosphodiester bond. The only direct sequence requirement in this region is a G:U wobble base pair formed between a U at the 3' end of the exon and a G located in the IGS at a phylogenetically conserved position at the exon/intron boundary (Michel & Westhof, 1990). The P1 helix is correctly positioned in the catalytic core by a number of tertiary interactions between the helix and conserved sequence elements within the ribozyme active site. These interactions include contacts from 2'-hydroxyls of the nucleotides in the P1 helix and a sequence specific tertiary contact to the exocyclic amine of the G:U wobble pair (Bevilacqua & Turner, 1991; Pyle & Cech, 1991; Strobel & Cech, 1993; Strobel & Cech, 1996; Strobel *et al*, 1998).

The *Tetrahymena* intron has been modified to generate a true ribozyme capable of multiple turnover (Zaug et al., 1986; Zaug *et al.*, 1988). The intron has been shortened at both the 5' and 3' ends. Deletion of 21 nucleotides from the 5' end removes unnecessary sequences from the P1 helix and allows binding and cleavage of an RNA substrate added in trans, and deletion of five nucleotides from the 3' end removes the 3' splice site and prevents the G at the 3' splice site from participating in the reaction. The resulting ribozyme is referred to as the L-21 Sca I ribozyme due to the removal of the 21 bases from the 5' sequence and its transcription from a plasmid template linearized by the Sca I enzyme. The L-21 Sca I ribozyme can cleave an exogenous RNA substrate *in vitro* in the presence of moderate concentrations of Mg^{2+}, and the enzymatic activity of this modified form of the intron has been well studied (Herschlag & Cech, 1990a; Herschlag & Cech, 1990b). Under multiple turnover conditions, the rate limiting step in the reaction is product release with a rate of .1 min^{-1}, consistent with the evolution of the full length intron to perform only a

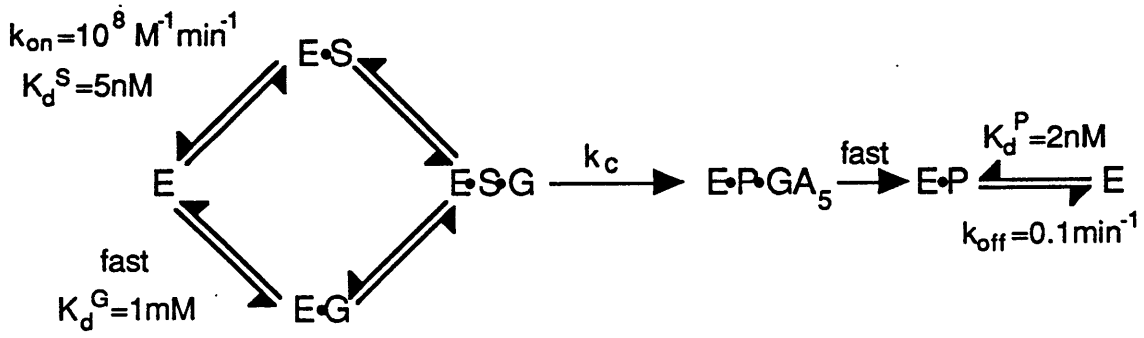


Figure 5 : Kinetic mechanism for the L-21 Sca I ribozyme. Adapted from Herschlag and Khosla (1994). E is the ribozyme, S is the substrate ($CCCUCUA_5$), GA_5 and P (CCCUC) are the reaction products, and G is the guanosine cofactor.

single self-splicing event. When substrate concentrations are subsaturating, the rate limiting step is substrate binding and the rate of chemistry is 350 min^{-1}. The k_{cat}/K_M for the ribozyme is 9 x 10^7 $M^{-1}min^{-1}$, while this is below the diffusional limit of 10^{11} $M^{-1}min^{-1}$ for collision of small molecules it is similar to the rate constants for helix formation between two oligonucleotides. Experiments monitoring binding using a fluorescently labeled substrate have shown that binding is a two step process involving simple helix formation followed by a conformational change that docks the substrate helix onto the catalytic core (Bevilacqua *et al.*, 1992). It is possible to monitor multiple steps in the cleavage reaction through varying the pH, GTP concentration, and concentration of substrate and enzyme, and this ability is useful in studying the formation of catalytically active ribozyme structure.

Structure

The ability to fold into a complex three dimensional structure is required for the *Tetrahymena* ribozyme to catalyze the splicing reaction. The structure of the ribozyme has been well studied and models of both the secondary structure and the tertiary structure have been proposed (Burke *et al.*, 1987; Lehnert et al., 1996; Michel & Westhof, 1990). All group I introns share a conserved catalytic core and are defined by the presence of conserved sequences and secondary structure motifs (Davies *et al.*, 1982; Michel *et al.*, 1982). A secondary structure representation for the *Tetrahymena* ribozyme that also depicts elements of the tertiary structure is shown in figure 6 and a model for the three dimensional structure is shown in figure 7. The catalytic core consists of two structural domains, the P4-P6 domain consisting of coaxially stacked helices P4, P5, and P6, and the P3-P7 domain consisting of coaxially stacked helices P7, P3, and P8. These two domains are aligned by two sets of nucleoside triple interactions called the triple helical scaffold. The nucleoside triples involve interactions between P4 and J6/7, and P6 and J3/4 and are supported by mutational analysis and covariations in phylogeny (Michel *et al.*, 1990). The P4-P6 and P3-P7 domains form a cleft that binds the P1 substrate helix, and this binding is stabilized by contacts between 2-hydroxyl groups from the P1 helix and the P4 helix. The

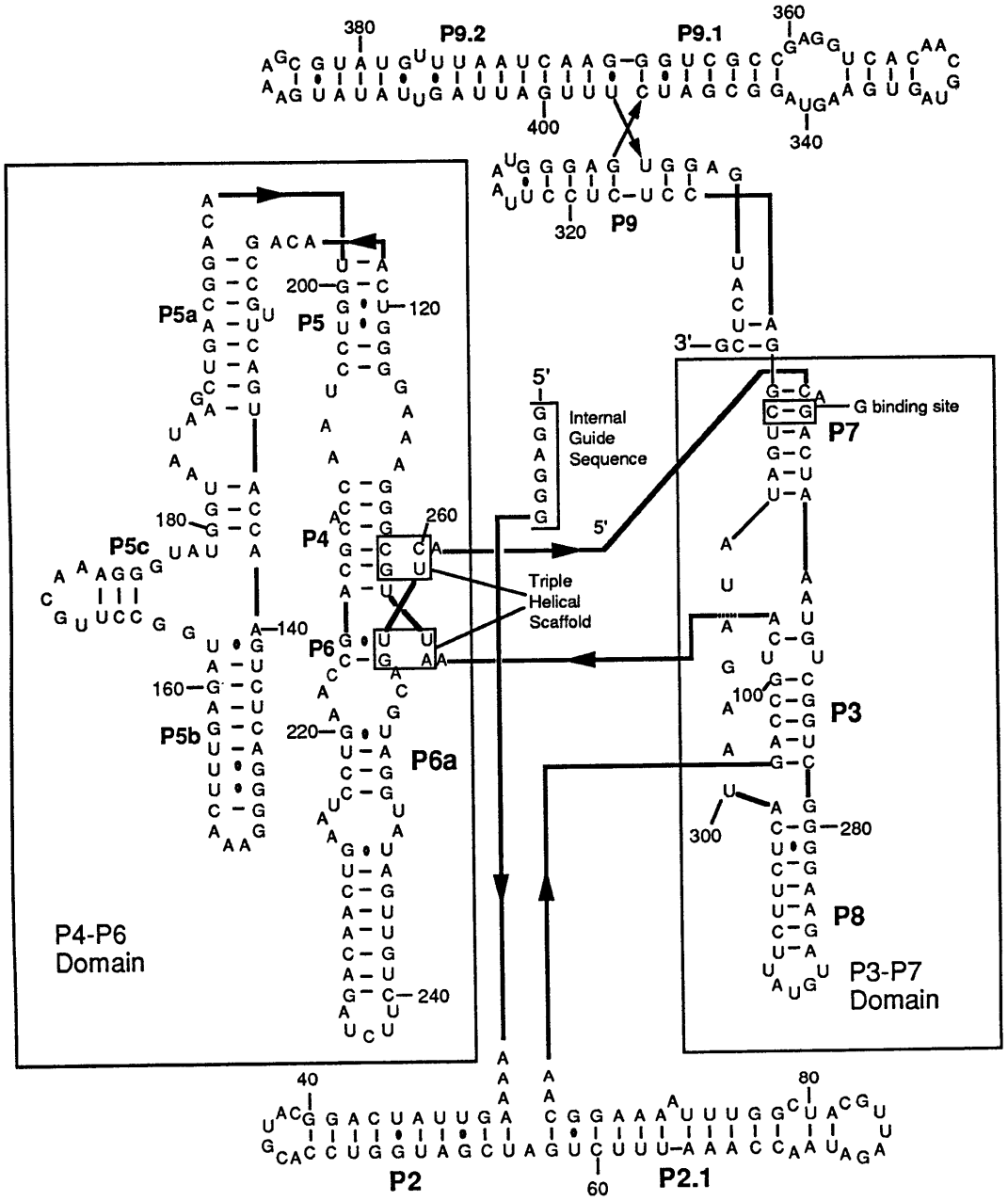


Figure 6 : The secondary structure of the *Tetrhaymena* ribozyme. The major folding domains and important tertiary contacts are boxed and labeled.

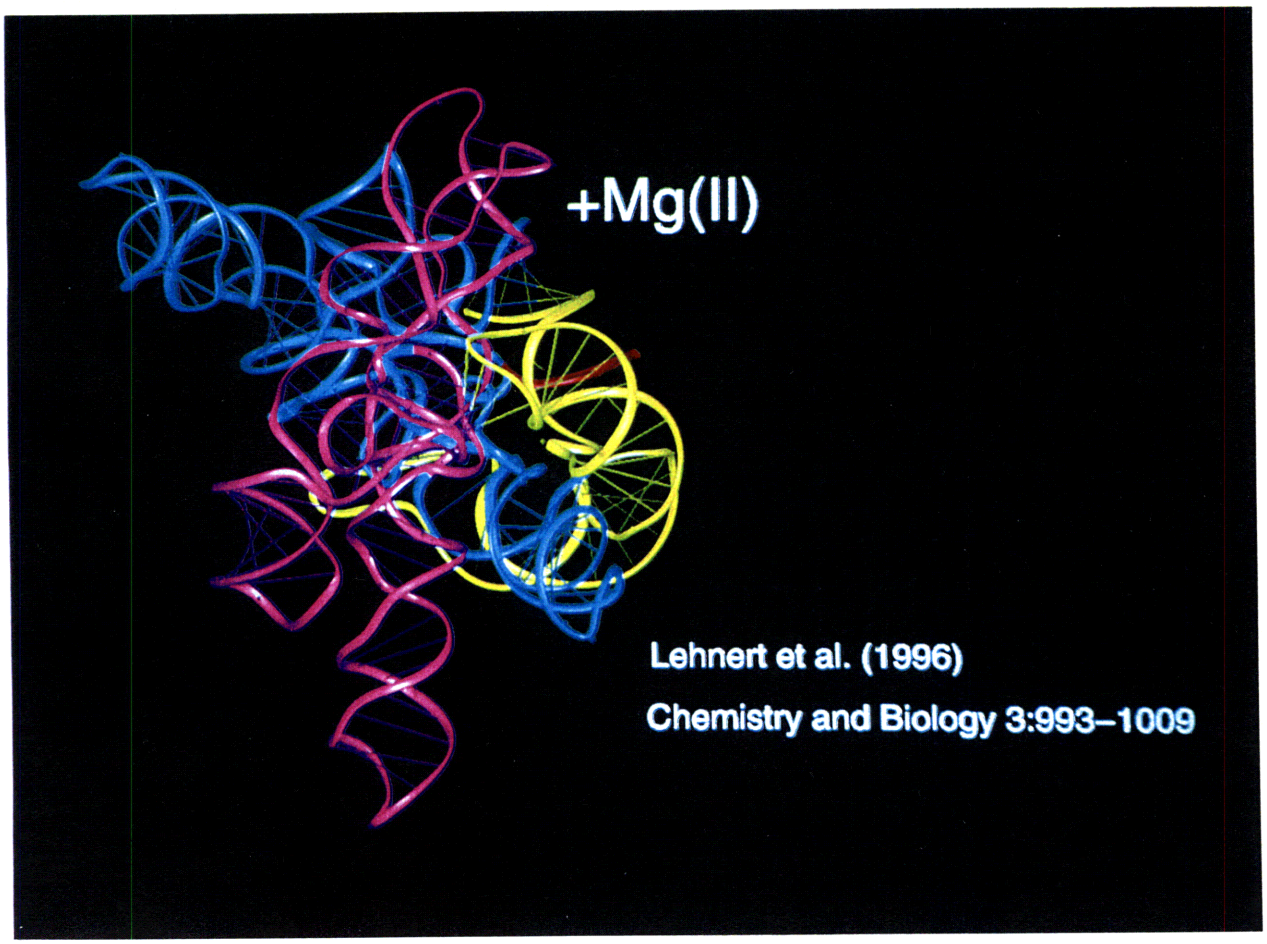


Figure 7: Proposed model for the three dimensional structure of the *Tetrahymena* ribozyme. The P4-P6 domain is shown in magenta, the P3-P7 and P9 domains are shown in blue, and the P2-P2.1 domain is shown in yellow.

ribozyme is also stabilized by a number of peripheral domains, P5abc, P2/2.1, and P9. P5abc contacts and stabilizes the P4/P6 domain. P2/2.1 consists of coaxially stacked P2 and P2.1 helices and wraps around and stabilizes the core. Additionally the loop of P2 forms tertiary contacts with the loop of P5c, while the loop of P2.1 contacts the loop of P9.1a.(Lehnert et al., 1996) Domain P9 is composed of helices P9, P9.1, and P9.2 and is thought to lock the P3-P7 core domain into place.

The P4-P6 domain (including the P5abc domain) has been shown to fold in isolation (Murphy & Cech, 1993; Murphy & Cech, 1994), and when the other structural domains are added in trans catalytic activity can be reconstituted. Recently a crystal structure of this isolated domain was solved (Cate et al., 1996a; Cate *et al.*, 1996b) and is shown in figure 10. The crystal structure gives much information about both the tertiary structure of the *Tetrahymena* ribozyme, but also about common structural motifs in RNA folding such as tetraloop-receptor contacts, backbone-backbone contacts and Mg^{2+} binding sites. The P4-P6 domain structure consists of two helical regions packed side by side. P6b, P6a, P6, P4 and P5 form a straight column followed by a 150° bend facilitated by the J5/5a internal loop and allowing the P5abc three-way junction to pack against the P4 and P6 helices. P5a and P5b are stacked causing P5c to protrude from the stacked structure. There are two tertiary interactions that act as clamps to hold P5abc and P4-P6 together. The first of these is an A rich bulge in P5a that is conserved and makes structural interactions critical to the proper folding of the domain. In the crystal structure, this bulge makes a corkscrew turn that flips out the bulge bases allowing them to interact with both P4 and the three-helix junction. Phosphates in the backbone are very close together and require the binding of at least two and as many as five magnesium ions. The second tertiary interaction is between the P5b GAAA tetraloop and an eleven nucleotide receptor in J6a/6b. A similar tetraloop/receptor interaction was found in the crystal structure of the hammerhead ribozyme (Pley *et al.*, 1994a). The loop docks in the minor groove of the P6a helix, and the three adenine bases are stacked on the bases on the 5' side of the tetraloop receptor helix forming an A-platform motif. A side-by-side

configuration of the adenosines results in a kink in the ribose backbone that opens the minor groove of the tetraloop receptor allowing the stacking to occur. The close packing that these two tertiary contacts mediate emphasizes the importance of magnesium ions in charge neutralization. While the crystal structure gives a number of new insights into RNA structure, there are still many questions about how these structures are formed. These questions must be addressed by experiments mapping out the folding pathways of RNA molecules.

Folding

The folding of RNA into complex three dimensional structures requires the presence of magnesium ions. The magnesium induced folding pathway of the *Tetrahymena* ribozyme is the pathway most often studied. Initializing folding by adding magnesium can be seen as analogous to the process of initializing protein folding by decreasing the concentration of denaturants. There are folding steps prior to the addition of magnesium, namely the formation of secondary structure, but formation of secondary structure is rapid (msec-μsec timescales (Cole & Crothers, 1972)) and has been studied in other simpler RNA systems such as tRNA and short RNA helices (Pörschke & Eigen, 1971). The unfolding of the *Tetrahymena* ribozyme has been analyzed using thermal unfolding monitored by UV absorption and chemical mapping (Banerjee *et al.*, 1993). These studies identified two major transitions in the thermal unfolding pathway and used chemical mapping to identify a high temperature transition as the unfolding of tertiary structure, and a lower temperature transition as the unfolding of secondary structure. This data suggested that the tertiary interactions are weaker than the secondary interactions and that the secondary structure dominates the free energy of folding.

The magnesium induced equilibrium folding pathway has also been analyzed. Free radicals produced by solvent-based Fe(II)-EDTA have been used to map regions of structure protected from solvent in the presence of varying concentrations of Mg^{2+} (Celander & Cech, 1991; Latham & Cech, 1989). The hydroxyl radicals interact with the ribose sugar of the RNA backbone resulting in sequence-independent strand scission, and the accessibility of the hydroxyl radical

probe is a good indication of solvent accessibility. The protection pattern of the *Tetrahymena* ribozyme supports the existence of higher order compact structure with an inside and an outside. Based on the Fe(II)-EDTA protection data an equilibrium folding pathway has been proposed. In the absence of magnesium, some secondary structure is formed but long range helices P3 and P7 are not. The P5abc region is the first to fold at low Mg^{2+} concentrations (0.6 mM) with the Hill coefficient of 3 (Murphy & Cech, 1993), followed by formation of the catalytic core at higher Mg^{2+} concentrations (0.9 mM) with a Hill coefficient of 8. The crystal structure of the P4-P6 domain has shown that the P5abc domain binds five magnesium ions, which correlates well with the Hill coefficient for this region. It has been suggested that the binding of these ions drives folding in much the same way that hydrophobic collapse drives protein folding (Cate et al., 1997).

The Fe(II)-EDTA assay for structure formation allowed the analysis of the importance of the peripheral domains in stabilizing the structure. The P4-P6 domain, including the P5abc subdomain, has been shown fold to its native structure in the absence of the rest of the ribozyme at normal magnesium concentrations (Murphy & Cech, 1993). Therefore, the P4-P6 domain is an independent folding domain, and RNA folding appears to have a modular organization much like proteins. The magnesium dependence of P3-P7 folding is slightly increased when the P9 domain is absent, indicating that the P9 domain plays a role in stabilizing this region of the core. The magnesium dependence of P4-P6 folding is unaffected by the absence of the P9 domain, but the P4-P6 domain is more destabilized by mutations in the P5abc region when the P9 domain is absent. Therefore, it appears that P9 also plays a minor role in stabilizing the P4-P6 domain as well.

The modular nature of the ribozyme is shown by experiments that divided the ribozyme into three domains that when reconstituted in trans can self-assemble into a catalytically active structure (Doudna & Cech, 1995). The three domains are held together mainly by tertiary interactions. Recovery of activity in the three piece system required the correct positioning of the bases involved in the three helix junction, highlighting the importance of the three helix junction in

properly aligning the P4-P6 and P3-P7 domains. The ability of the three piece ribozyme system to self-assemble led to another question about folding, do the three domains fold independently and then assemble through the tertiary interactions or is there an order to the folding process? Fe(II)-EDTA mapping of the three separate domains showed that only the P4-P6 domain could fold independently (Doherty & Doudna, 1997), P3-P7 showed no evidence of higher order folding unless P4-P6 was present. This led to the proposal that a folded P4-P6 domain is required to template folding of the rest of the ribozyme, and indicated that at least equilibrium folding proceeded in a hierarchical manner.

While the equilibrium folding pathway can give information about structures that may correspond to kinetic intermediates, it cannot define the kinetic folding pathway. To define a kinetic folding pathway folding rates must be measured directly. A kinetic oligonucleotide hybridization assay was developed in the Williamson laboratory to measure rates of RNA folding (Zarrinkar & Williamson, 1994). This technique takes advantage of changes in the accessibility of sequences to complementary DNA oligonucleotides after folding (Figure 8). Different regions of the RNA can be studied specifically and independently by using probes with the appropriate complementarity. While the accessibility to probe binding is mainly determined by secondary structure formation, tertiary structure formation can also be monitored indirectly if it is required to stabilize secondary structure. Folding is initiated by the addition of Mg^{2+}, and the fraction of RNA accessible to oligonucleotide binding is determined as a function of folding time by cleavage of the oligonucleotide-RNA complexes with RNase H. The reaction is in effect a rapid quench experiment that measures the fraction of RNA that remains unfolded at any given time after folding is initiated. To ensure that the fraction of RNA uncleaved accurately reports the fraction of RNA folded four kinetic conditions must be met: 1) the rates of probe hybridization and RNase H cleavage must be rapid compared to the folding rate, 2) sufficiently high probe concentrations must be used so that all accessible RNA is bound, 3) RNase H cleavage must be much faster than dissociation of the RNA-probe hybrid, and 4) unfolding under native conditions must be slow

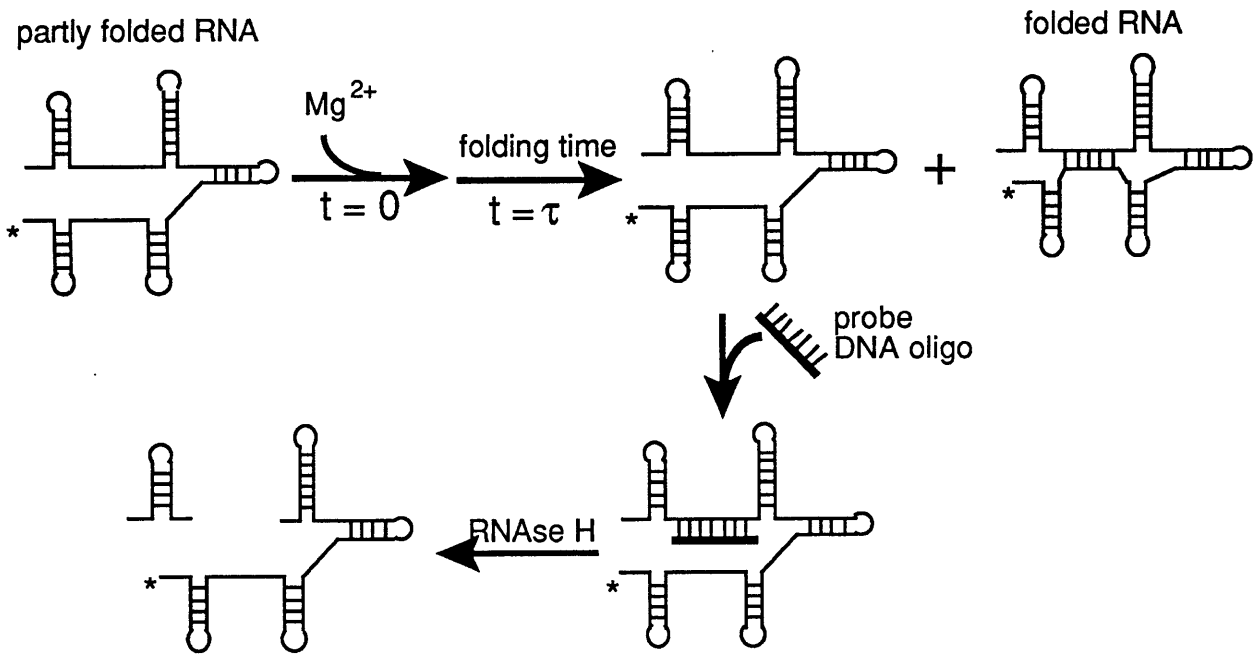


Figure 8 : The kinetic oligonucleotide hybridization folding assay.

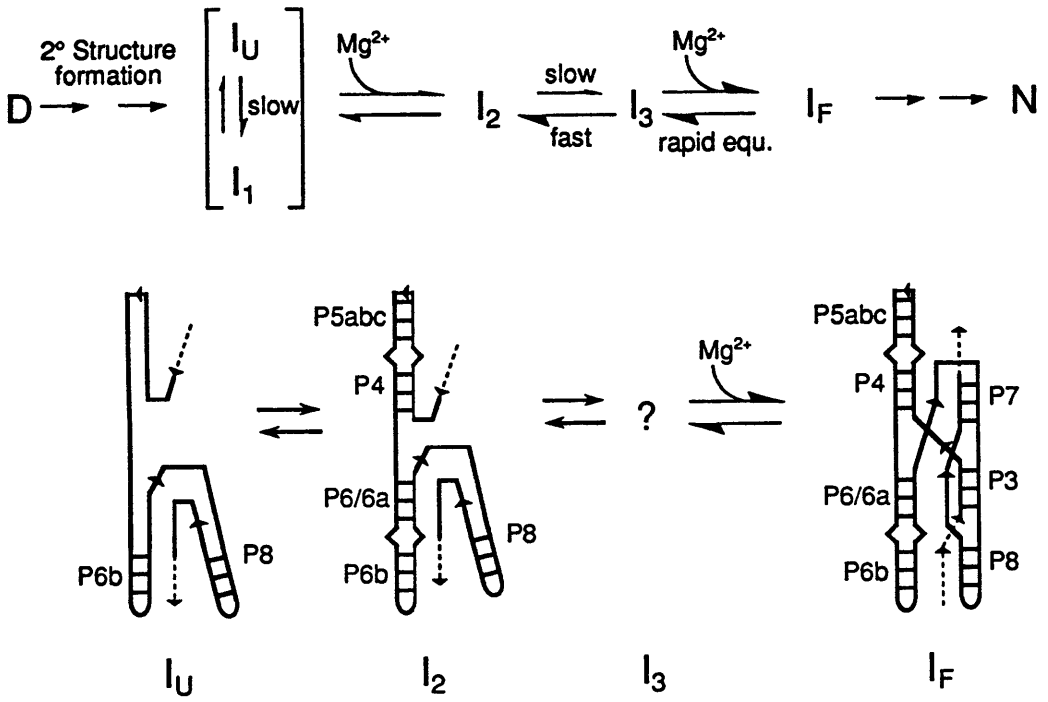


Figure 9 : Model for the kinetic folding pathway of the *Tetrahymena* ribozyme.

compared to folding. The oligonucleotide hybridization folding assay was developed to meet all of these conditions.

Based on data from the oligonucleotide hybridization assay a folding pathway for the *Tetrahymena* ribozyme was proposed and is shown in Figure 9. One limitation of this assay is that only steps that occur after Mg^{2+} addition and steps that result in the change of accessibility to the oligonucleotide probes can be measured. Therefore, the pathway starts with denatured RNA (D) and a number of steps leading to secondary structure are represented and lead to the formation of an intermediate I_U. When probes targeting the P4-P6 region were added at the same time as Mg^{2+} approximately sixty percent of the RNA molecules were inaccessible to probe binding and RNase H cleavage. If limited cleavage were due to similar rates for folding and probe binding, the extent of cleavage should increase with increasing concentrations of probe. The extent of cleavage was unaffected by increased concentrations of oligonucleotide probes. It appears that the population of unfolded molecules is heterogeneous with only a subpopulation of molecules accessible to probe binding. If probes targeting P4 or P6 are added before the addition of Mg^{2+}, the extent of cleavage increases to above eighty percent. Therefore, it was proposed that there is an equilibrium between two conformations, one inaccessible to probe binding that can fold rapidly, the other that must fold slowly. These two populations are represented by I_U and I_1 on the folding pathway, since they may not represent a productive step on the pathway they are in brackets. While the amplitude of the folding phase was too small to measure accurately the magnesium dependence of the rate of P4/P6 folding, magnesium is required for folding and the requirement is indicated on the pathway. There is an intermediate (I_2) on the pathway in which P4-P6 is folded and P3-P7 is not. This intermediate is evident because the rate of folding for P3 and P7 (0.7 min^{-1}) is slow compared to the rate for P4-P6 folding. P3 and P7 both fold with the same rate, but additional experiments were needed to determine if P3 and P7 folded cooperatively. Mutations in P3 that increased the magnesium requirement for folding also increased the magnesium requirement for P7, and mutations in P7 that increased the magnesium requirement for P7 folding also increased it for P3

folding. This indicates that P3 and P7 fold in a interdependent manner. Magnesium is required for P3-P7 formation, but the folding rate is independent of magnesium concentration. This suggests the presence of a slow unimolecular rearrangement leading to the formation of another intermediate on the folding pathway (I_3) and followed by a rapid magnesium binding step. Mutations in the triple helical scaffold slow the folding of P3-P7 indicating that the triple helical scaffold is formed during the slow step. The slow unimolecular rearrangement leading to P3-P7 formation was determined to be the slowest step on the folding pathway by also measuring the rate of formation of the catalytically active structure. This assay measures the slowest step on the folding pathway and the rate for formation of the catalytically active structure was 0.6 min^{-1}, within error of the rate of P3-P7 formation.

The proposed folding pathway has been supported by a number of subsequent experiments. Ultra-violet inducible crosslinking experiments have been used to explore the folding of the *Tetrahymena* ribozyme (Downs & Cech, 1990; Downs & Cech, 1996). When RNA is subjected to short-wave UV light (λ_{max} = 254 nm) aromatic electrons of the heterocyclic bases are excited and can form new chemical bonds with other nucleotides. If crosslinks are formed between bases far apart in the linear sequence, they can report on the formation of tertiary structure. Such crosslinking was shown to form three primary crosslinked adducts in the folded structure of the ribozyme; between the P2 and P2.1 helices showing stacking of the two helices, at the bend in J5/5a, and in the triple helical scaffold that implies the presence of non-native structure. A folding assay was used to monitor the rate of folding based on the rate of adduct formation after initiation of folding by addition of magnesium. Formation of the bend in J5/5a occurred first, followed by formation of the non-native structure in the triple helical scaffold. This non-native structure eventually decreased, with approximately the same rate as formation of the P2/P2.1 crosslink. All folding was completed with a rate of 0.82 min^{-1}, consistent with the rate of the slowest step measured by the oligonucleotide hybridization assay.

A more detailed understanding of the folding rates for P4-P6 has come from time-resolved hydroxyl radical footprinting experiments (Sclavi *et al.*, 1998; Scott et al., 1995). These experiments used the same hydroxyl radical mechanism of cleavage of the phophodiester backbone as in Fe(II)-EDTA footprinting, however the hydroxyl radicals are generated using high flux x-ray beams. The use of x-ray generated hydroxyl radicals allows footprinting to be accomplished with millisecond time resolution. The P5abc region, specifically the A-rich bulge and P5b tetraloop-receptor tertiary interactions, was the first to form with rates of 1.5 s^{-1}, the rest of the P4-P6 domain forms next with rates of 1 s^{-1}. This is consistent with the proposal that Mg^{2+} induced folding of the P5abc region nucleates folding of the rest of the ribozyme. Protection in the P2-P2.1 and the P9 domains occurs next with a rate of 0.3 sec^{-1}, this leads to an intermediate with folded P4-P6 domain stabilized by peripheral domains P2 and P9 but a largely disorganized P3-P7 domain. The P3-P7 domain folds with a slow rate of 1 min^{-1} and leads to further protections on the face of P4 and P6 helices where P3-P7 packs.

Once the ribozyme structure is folded, the substrate must bind to the internal guide sequence and form the P1 helix and the P1 helix must dock into the catalytic core. The rates of these steps have been addressed using fluorescently labeled substrate molecules to monitor the binding of substrate (Bevilacqua et al., 1992). The binding was found to be a two step process, with simple helix formation occurring first with a rate of 3.9 x 10^6 $M^{-1}s^{-1}$ that is the rate expected for simple helix formation. Once the duplex is formed, a conformational change leading to docking of the P1 helix by tertiary contacts (see Structure section) occurs with a rate of 2.5 s^{-1}.

These data yield an overall picture of the steps along the folding pathway of the *Tetrahymena* ribozyme, and the rates at which these steps occur. The folding is a multistep process that occurs in a hierarchical manner with P4-P6 templating formation of P3-P7 and requiring peripheral folding domains to stabilize the overall fold. There is at least one slow step on the folding pathway, and perhaps one non-native intermediate. However, the details of what makes the folding of P3-P7 slow were still in question. Small RNA molecules such as tRNA are

known to fold on the millisecond timescale. One domain of the *Tetrahymena* ribozyme can fold on the millisecond timescale, why doesn't the entire structure fold this fast?

Mutations that accelerate folding can provide insight into the mechanistic basis of slow folding steps (Rothwarf & Scheraga, 1996). Therefore, an *in vitro* selection procedure was developed to identify mutant *Tetrahymena* ribozymes in which the slow P3-P7 folding step is accelerated (Treiber *et al.*, 1998). Ribozymes that folded rapidly after Mg^{2+} addition were selected from a pool of RNAs containing an average of four mutations per molecule. Slow folding RNAs were selectively depleted from the pool by kinetic oligonucleotide hybridization using probes targeting P3 and P7. A step was included in each cycle of selection to ensure that fast-folding mutants formed an intact catalytic core. After nine rounds, the folding rate of the pool had increased by a factor of four over the initial pool and by a factor of two over wild type. The pool was then analyzed and four point mutations that all localized to the P5abc region of the P4-P6 domain were found to accelerate P3-P7 folding: A183U, A171G , +G174 (a mutant in which a G is inserted at position 174), and U167C. Each of these mutations can disrupt tertiary interactions that are important for P4-P6 formation. A183 is located in the highly conserved A-rich bulge and makes two hydrogen bonds that bridge the helical stacks in P4-P6 (Cate et al., 1996a). A171 is part of an adenosine platform motif in the loop of P5c (Cate et al., 1996a; Cate et al., 1996b) and is proposed to be involved in a pairing between L5c and the loop of P2 (Lehnert et al., 1996). In the +G174 mutant, addition of an extra guanosine could disrupt the base pairing or pairing register of P5c, and the U167C mutation disrupts basepairing in P5c. The localization of these mutations to one structural subdomain suggested a common mechanism of action. The mutations do not greatly affect either the catalysis or the stability of P3-P7. Although P5abc makes no direct contacts with P3-P7 in the three dimensional structure model, and the P4-P6 domain aquires its native structure prior to P3-P7 formation, positions in P5abc influence the folding rate of P3-P7.

The P5abc mutations are all found in positions that could disrupt the magnesium core of P4-P6 that is believed to nucleate P4-P6 folding. Since atom changes that disrupt the Mg^{2+}

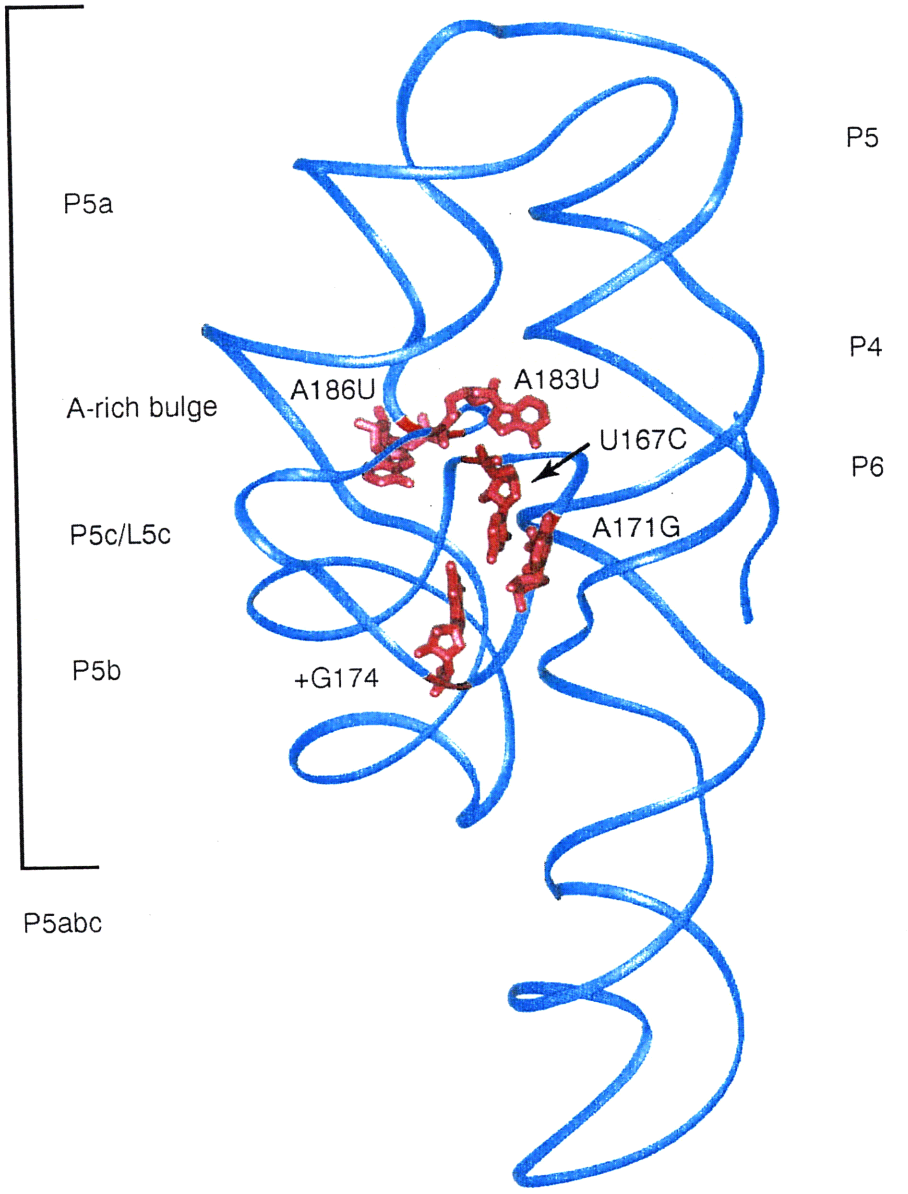


Figure 10: The structure of the P4-P6 domain. Structural features are labeled. The mutations that accelerate P3-P7 folding are shown in red.

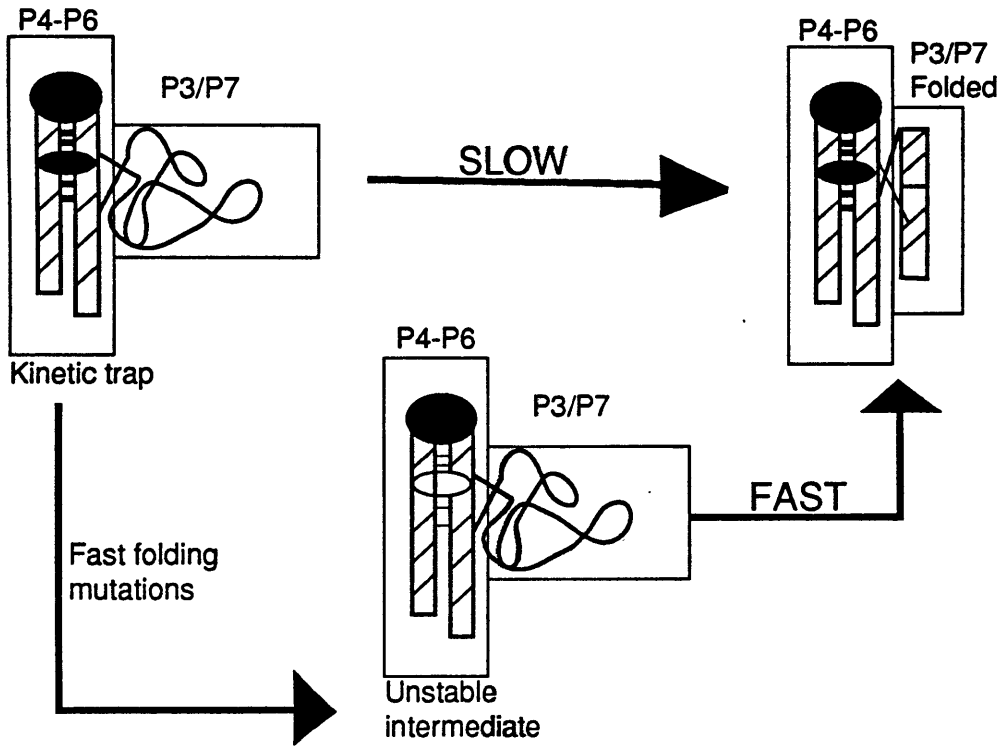


Figure 11: Model of a kinetic trap that slows P3-P7 folding. A kinetic trap stabilized by native structure in the P4-P6 domain slows formation of P3-P7. The mutations destabilize the P4-P6 domain and allows rapid formation of the P3-P7 domain.

binding sites destabilize the entire P4-P6 domain, destabilization of P4-P6 in I_2 may accelerate P3-P7 formation. In support of this theory, the A186U mutation, a known mutation that destabilizes P4-P6 formation, was found to accelerate P3-P7 folding as well. Since destabilization of native P4-P6 accelerates P3-P7 formation, it was proposed that I_2 is a kinetic trap. This trap contains both native (P4-P6) and non-native (P3-P7) interactions; it also corresponds well with the non-native structure found in the UV crosslinking experiments. Each of these structural features may restrict the conformational flexibility of I_2 and slow P3-P7 formation. Therefore, it appears that native structure formation in P4-P6 slows the formation of P3-P7.

2. Probing kinetic folding barriers by analyzing mutant ribozymes with accelerated P3-P7 folding.

Understanding the nature of the kinetic barriers to folding can help to explain the fundamental steps required for RNA folding. The position of the mutations in the P5abc region lead to the hypothesis that a kinetic trap stabilized by native contacts slows P3-P7 folding in the *Tetrahymena* ribozyme, but additional evidence is required to confirm the presence of a kinetic trap. The effects of mutations, temperature, magnesium, and denaturants on P3 folding in the wild type ribozyme were analyzed to explore the kinetic folding pathway and to determine the molecular nature of the kinetic barrier to P3 folding.

The effect of the selected mutations on the kinetic folding of the P4-P6 and P3-P7 domains.

The folding of the P4-P6 and P3-P7 domains was measured for the wild type and four of the fast folding mutant ribozymes (A183U, A171G , +G174 (a mutant in which a G is inserted at position 174), and U167C) using the kinetic oligonucleotide hybridization assay. This assay measures the time dependence of changes in the accessibility of the ribozyme to complementary oligonucleotide probes. Probes targeting the P6 helix report on the folding of the P4-P6 domain, while probes targeting the P3 helix report on the folding of the P3-P7 domain. For the kinetic hybridization assay to accurately reflect folding, the rates of probe hybridization must be rapid compared to the folding rate and all accessible RNA must be bound by the probe and cleaved by RNase H. Since the mutations may affect probe binding, control experiments testing the extent of cleavage of unfolded mutant RNA at a series of concentrations of oligonucleotide probe were performed for probes targeting P3 and P6 (Figure 12B and 13A). These controls showed that the mutations do not affect probe binding, and that at concentrations between 20 and 100 μM of the P3 oligonucleotide greater than 90 percent of the RNA was cleaved. A higher concentration of probes

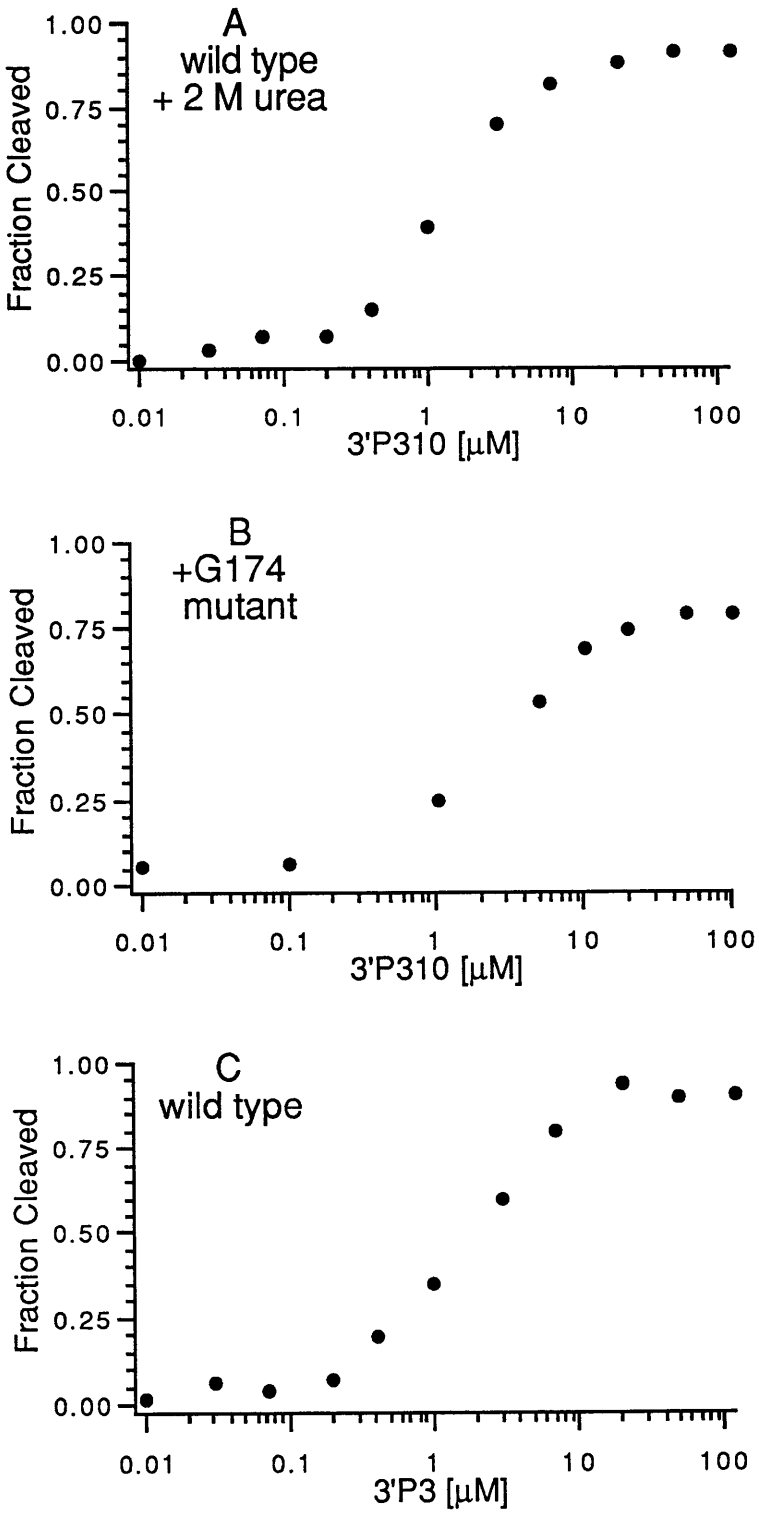


Figure 12: The concentration dependence of P3 oligonucleotide probe binding and RNase H cleavage. (A) wild type in the presence of 2 M urea (B) the +G174 mutant (C) wild type

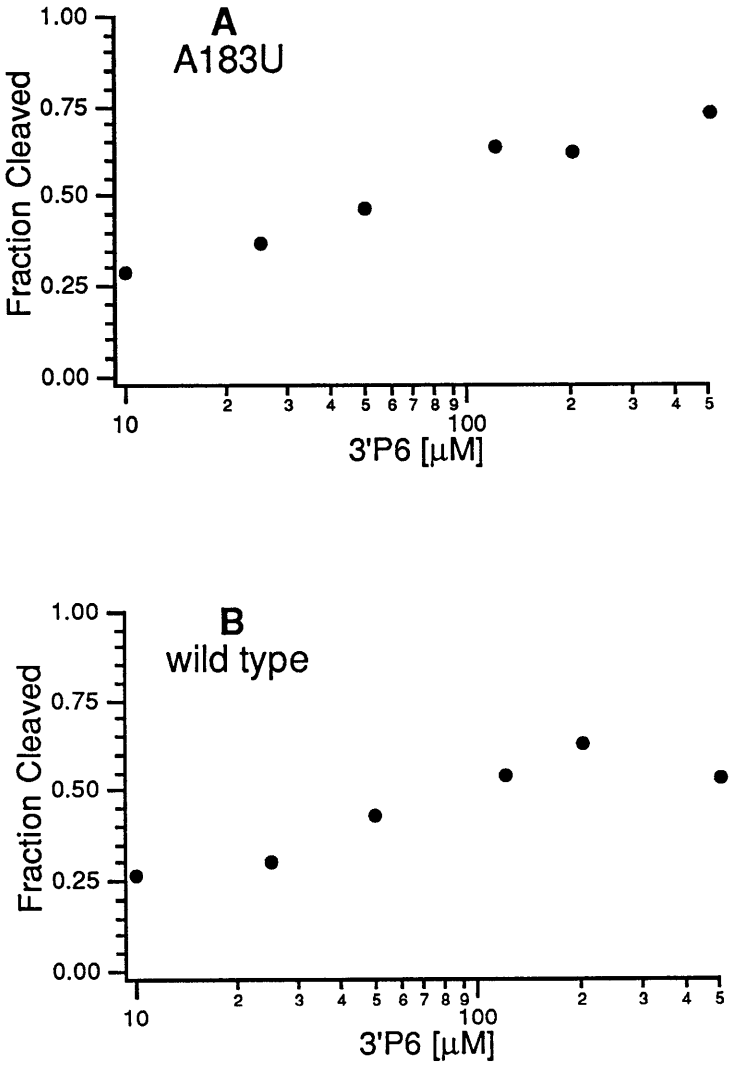


Figure 13: The concentration dependence of P6 oligonucleotide probe binding and RNase H cleavage. (A) the A183U mutant (B) wild type

targeting P6 (200 μM) was required for maximal cleavage, and only 60 percent of the RNA was accesible to cleavage in both the wild type and mutant ribozymes.

Once it was determined that the mutations do not perturb the folding assay, the folding rates of the P3-P7 and P4-P6 domains were measured. The folding of the P3-P7 domain is accelerated by the mutations (Figure 14A), while the folding of the P4-P6 domain appears unchanged by the mutations (Figure 14B). It appears that the mutations affect only P3-P7 folding, therefore folding experiments focused on the folding of the P3-P7 domain.

The mutations that accelerate P3-P7 folding are clustered in the P5abc subdomain of the ribozyme, because of this clustering it was hypothesized that they share a common mechanism for accelerating folding. The kinetic traces for P3 folding of the four fast folding mutants are shown in Figure 15. While the P3 folding rates were very similar for the four mutants, the A183U and A171G mutants appear to fold with biphasic kinetics. The amplitude of the slow phase varied in different folding experiments from ten to thirty percent. To determine if the data were best fit by single or double exponential curves the data from at least three experiments was averaged and the standard deviation of the fraction folded at each time point was determined. Linear least squares analysis was used to fit the average data to both single exponential and double exponential kinetic curves. The standard deviation for each time point was used as a weighting factor to determine the goodness of fit and to obtain errors for the amplitudes and rates measured. Based on this procedure, the U167C and +G174 mutants were determined to be best fit by a single exponential, while the A171G and A183U mutants were best fit by double exponentials. It appears that in the A171G and the A183U mutant ribozymes there are two folding populations, one which folds rapidly ($\cong$ 10 min^{-1}) and a second which folds slowly ($\cong$ 0.7 min^{-1}). While the slow rate is very similar to the rate for wild type P3 folding it does not appear that the slow population is following the wild type folding pathway (see temperature dependence section), it may instead reflect a fraction of the molecules in which the mutations stabilize a misfolded structure.

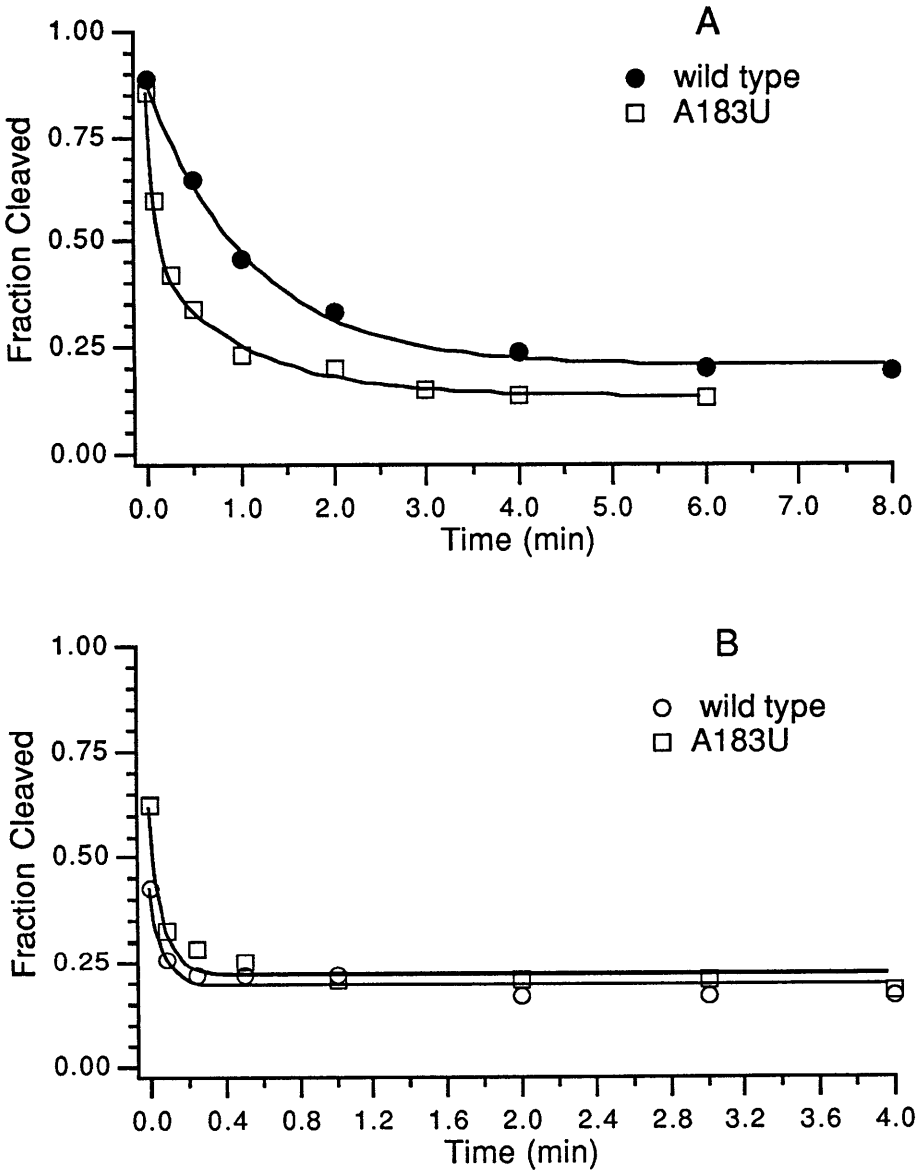


Figure 14. Folding of the P3-P7 domain is accelerated by the fast folding mutations (A) while folding of the P4-P6 domain is unaffected (B).

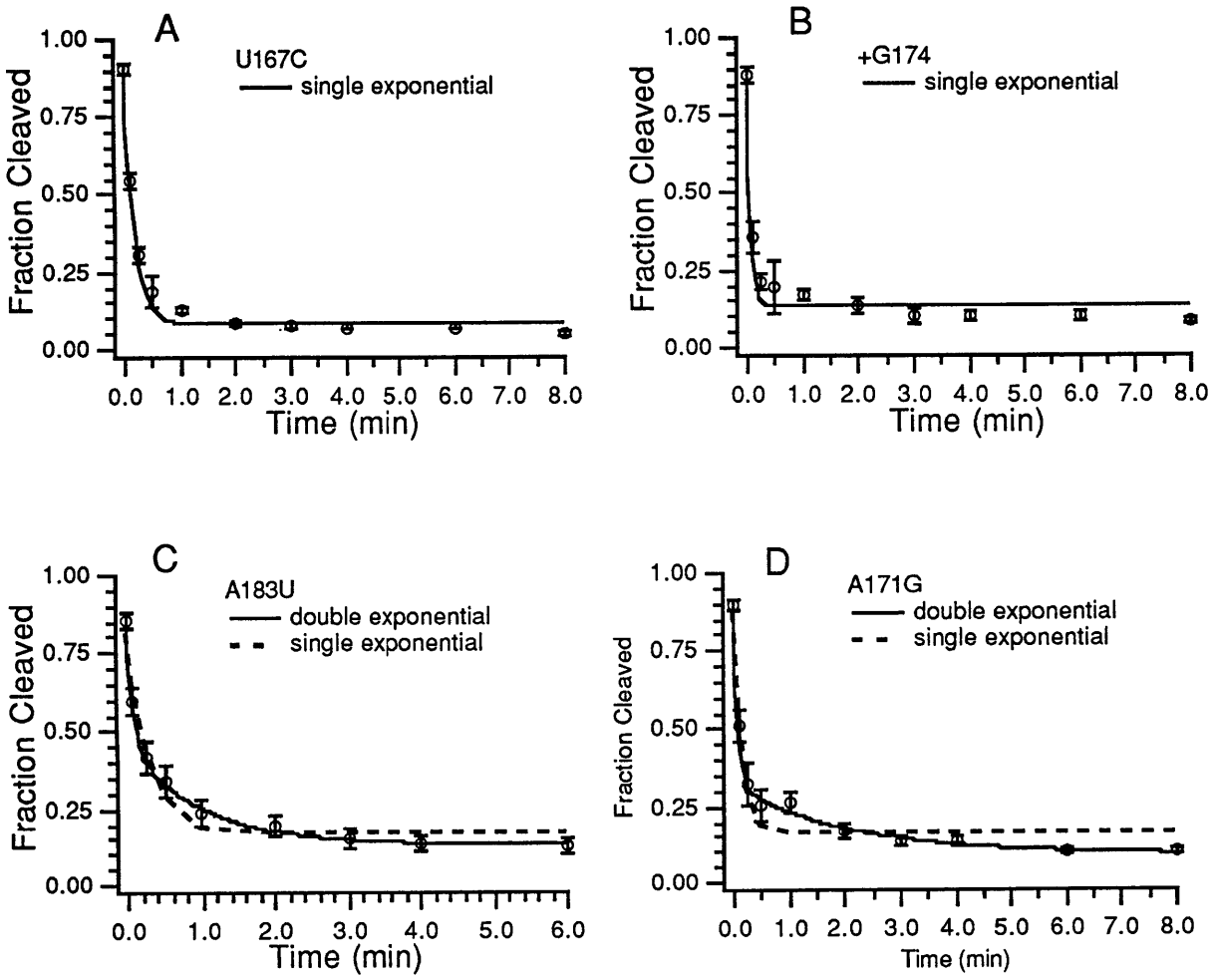


Figure 15. Kinetic traces for P3 folding of four fast folding mutants.

The temperature dependence of folding.

We wished to determine how the fast folding mutations affect the kinetic barriers to P3 folding. The temperature dependence of a folding rate can give information about the height of an activation energy barrier by giving values for the activation enthalpy and entropy of folding (Oliveberg *et al.*, 1995). Activation enthalpy can be correlated to energy required to break bonds or disrupt structure to achieve the transition state (Tinoco *et al.*, 1995). A high activation enthalpy and a low activation entropy are characteristic of unfolding in proteins, and would be indicative of the presence of a kinetic trap in a particular folding step. Therefore, the rate of wild type P3 folding at various temperatures was measured and compared with the rates for the selected fast folding mutants.

To measure magnesium induced folding the RNA must first be annealed to form the correct secondary stucture. When folding experiments were performed at low temperatures ($\leq 27°C$) the annealing conditions were found to have an effect on the folding rates of both P3 and P6 (fig. 16). If the concentration of Na^+ was lower than 2 mM the rate of both P3 and P6 folding slowed, and the extent of cleavage of P6 at the zero timepoint increased. At 37°C, the concentration of sodium in the annealing mixture did not affect the rate of folding. The change in the rate most likely reflects misfolded secondary structure at low sodium, but as we wished to explore only the magnesium induced folding pathway annealing was always performed in the presence of 2 mM Na^+.

When the annealing buffer contained 2 mM Na^+ the rate of P6 folding was independent of temperature. However, the amplitude change for folding was very small and the majority of the change occurs in the dead time of the folding experiment (5 seconds). There may be a temperature dependence that cannot be measured by the oligonucleotide hybridization assay.

The temperature dependence of the P3 folding rates for the wild type and mutant ribozymes is shown in Figure 17. The data was analyzed using Eyring transition state theory, and plots of

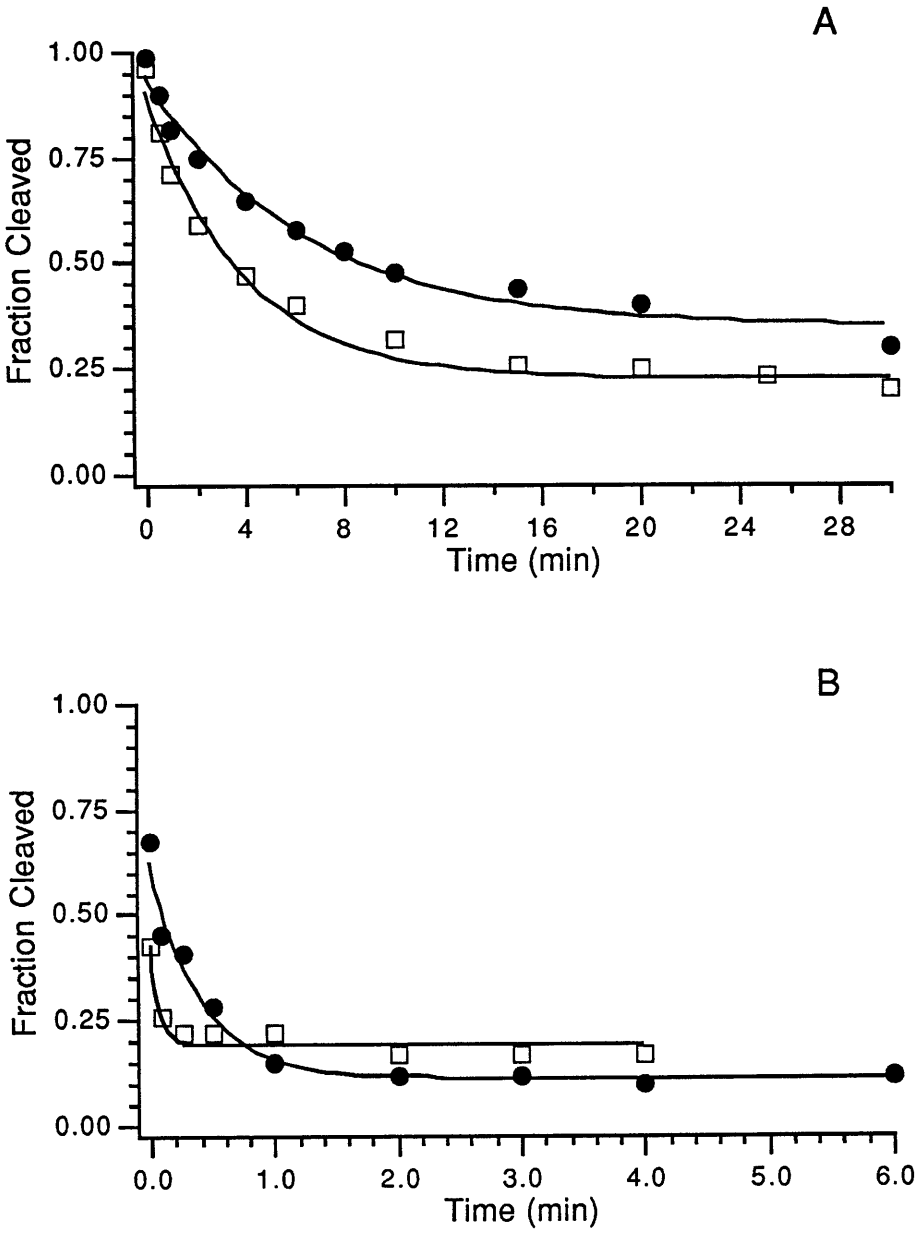


Figure 16. The effect of annealing conditions on the folding rate of P3 and P6.
A. P3 folding (●) annealing in 0.2 mM Na^+, (❑) annealing in 2 mM Na^+
B. P6 folding (●) annealing in 0.2 mM Na^+, (❑) annealing in 2 mM Na^+

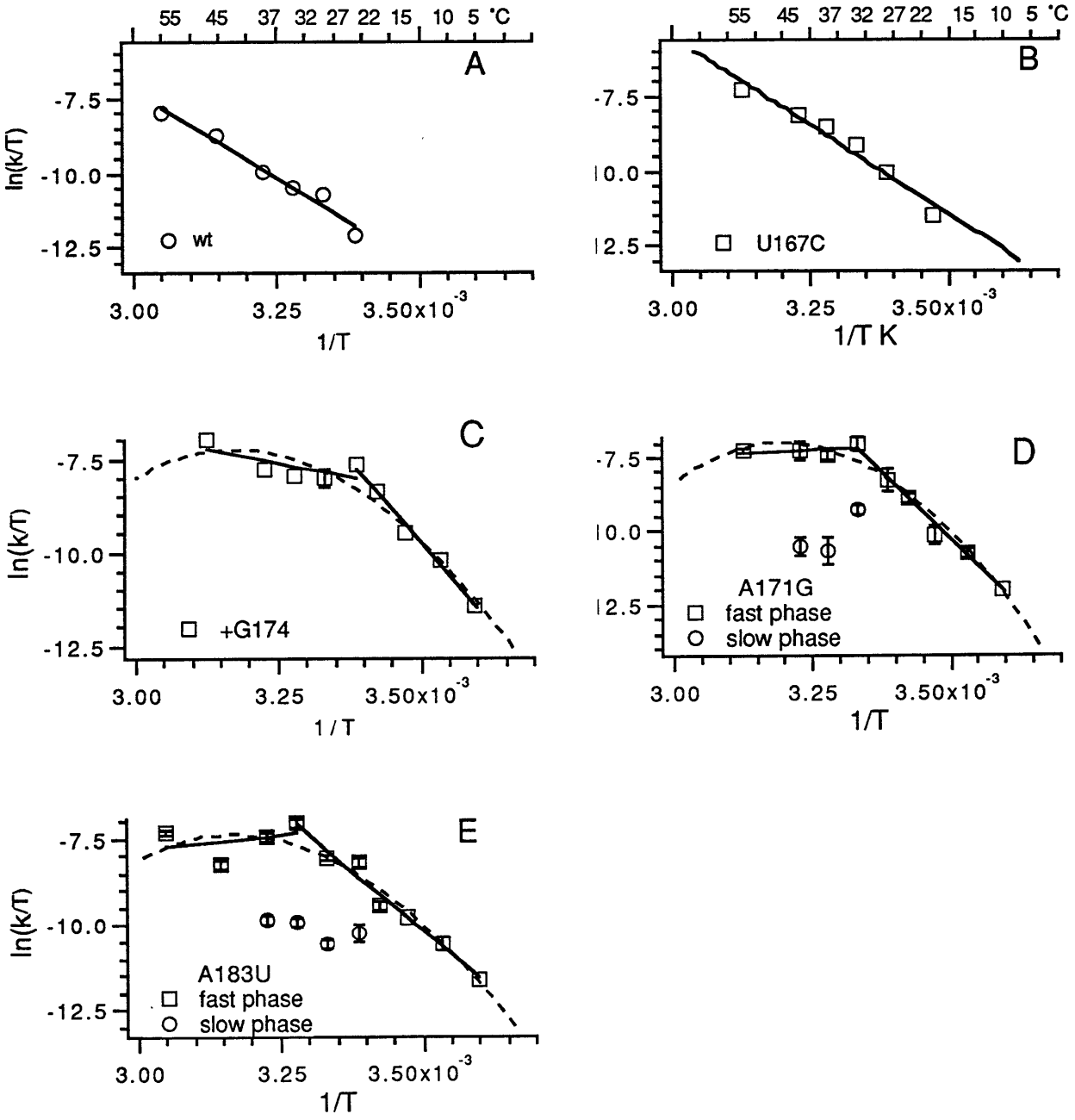


Figure 17. Eyring plots of the temperature dependence of P3 folding. Error bars are less than or equal to the size of the marker unless otherwise shown. Each point represents at least two and as many as five experiments. Solid lines represent fits to the Eyring equation, while dashed lines represent fits to the Eyring equation modified to account for a change in the heat capacity.

$\ln(k_{obs}/T)$ vs $1/T$ were fit to equation 1

$$\ln(k_{obs}/T) = \ln(k_B/h) - \Delta H^{\ddagger}/RT + \Delta S^{\ddagger}/R \quad (1)$$

k_{obs}	=	observed folding rate	T	=	Temperature, °K
k_B	=	Boltzman constant	h	=	Planks constant
$\Delta H^{\ddagger}$	=	activation enthalpy	$\Delta S^{\ddagger}$	=	activation entropy
R	=	gas constant			

to obtain values for the activation enthalpy and entropy for P3 folding (table 1). For the mutants with biphasic folding only the fast folding rates were fit to the Eyring equation because only the fast folding population gives information on how the mutants evade the kinetic trap.

The temperature dependence of the rate of P3 folding for wild type ribozyme is linear and has an activation enthalpy for P3 folding of 23±0.6 kcal/mol, which agrees with previously measured values (Downs & Cech, 1996; Zarrinkar & Williamson, 1994). This indicates that bonds must be broken, or that structure must be disrupted, to reach the transition state. The wild type has an activation entropy of 8.0±1.9 cal/mol-K, which indicates that achieving the transition state for P3 folding is favorable entropically. These values are consistent with the unfolding of an intermediate occurring during the P3 folding step.

Since mutations that disrupt native tertiary contacts accelerate P3-P7 formation (Treiber et al., 1998) , we hypothesized that the large $\Delta H^{\ddagger}$ for wild type P3 folding was due to the need to disrupt these interactions to reach the transition state. We expected that the mutant ribozymes would have a reduced activation enthalpy for P3 folding since these contacts were already disrupted or destabilized. Three of the four mutations (+G174, A183U, and A171G) decrease the activation enthalpy for P3 folding at the temperature at which the mutants were identified (37°C). The Erying plots for the A183U, A171G, and +G174 mutant ribozymes have many features in common. At temperatures above 22°C, the activation enthalpy is markedly reduced compared to the activation enthalpy of the wild type ribozyme (fig. 17 C,D, and E and Table 1). This indicates that less structure needs to be disrupted during P3 folding in the mutants than in the wild type.

Table 1. Activation enthalpy and entropy of ribozyme folding.

		$\Delta G^{\ddagger}$	$\Delta H^{\ddagger}$	$\Delta S^{\ddagger}$
ribozyme	(temp)	kcal/mol	kcal/mol	cal/mol-K
wild type	(37°C)	20.5±0.8	23.0±0.6	8.0±1.9
wt + 2M urea	(37°C)	19.8±3.2	15.0±2.1	-15.6±8.0
U167C	(37°C)	19.6±2.0	23.5±1.4	12.7±4.8
A171G ≥ 22°C	(37°C)	19.9±2.0	-1.3±1.4	- 65.6±4.4
≤ 22°C	(10°C)	19.4±2.5	36.7±1.8	61.4±6.1
+G174 ≥ 22°C	(37°C)	19.2±2.7	5.6±1.9	-43.9±6.0
≤ 22°C	(10°C)	19.1±2.5	35.4±1.8	57.6±6.2
A183U ≥ 22°C	(37°C)	19.0±1.5	-3.5±1.1	-73.0±3.5
≤ 22°C	(10°C)	19.2±1.2	27.7±0.9	29.8±3.0
A183U +2M urea				
fast phase	(37°C)	19.1±1.1	-3.5±1.1	-73.0±3.5
slow phase	(37°C)	19.9±2.3	50.0±1.6	97.2±5.4

The A183U and the A171G mutants both folded with biphasic kinetics at temperatures between 37°C and 22°C. At lower temperatures folding occurred in a single phase but only 50-60% of the population folded to completion, perhaps indicating that the slow folding population does not fold at very low temperatures. While the rates of the slow phases at 37°C were similar to that of the wild type, the temperature dependence of the slow phase is flat while that of the wild type is steep, suggesting that the two steps monitor different folding events.

In the Erying plots for +G174, A183U, and A171G there is a break in the temperature dependence at 22°C. Kinks or curvature in these plots are caused by a change in the activation enthalpy with temperature, which may be due to a difference in the heat capacity between the ground and transition states, temperature-induced changes in the ground state, a change in the rate limiting step, or some combination of these three mechanisms (Oliveberg et al., 1995). The Eyring equation can be modified to account for a heat capacity change during folding as shown in equation 2.

$$\ln(k_{obs}/T) = \ln(k_B/h) - \Delta H^{\ddagger}(T_o)/RT + \Delta S^{\ddagger}(T_o)/R - \Delta Cp^{\ddagger}(T-T_o)/RT + \Delta Cp^{\ddagger}\ln(T/T_o)/R \qquad (2)$$

k_{obs}	=	observed folding rate	T	=	Temperature, °K
T_o	=	reference temperature	k_B	=	Boltzman constant
h	=	Planks constant	$\Delta H^{\ddagger}$	=	activation enthalpy
$\Delta S^{\ddagger}$	=	activation entropy	R	=	gas constant

The plots were also fit to the modified equation and the fits are shown in Figure 17 as dashed lines. Since the experimental temperature range is rather small, the fits of the data to theoretical curves are not sufficient to conclude that the breaks in the plots are caused by a change in the heat capacity. The fits give a $\Delta Cp^{\ddagger}$ for P3 folding in all three of the mutant ribozymes of approximately 1000 cal/mol-K, which is comparable to that found for folding of fairly large proteins (Tan et al., 1996). It is appears that the break in the data is somewhat sharper than the fit to equation 2, and at higher temperatures where the fit predicts a decreased rate of folding, the rate of folding is actually still increasing. At temperatures above 55°C the rate is faster than can accurately be measured by

our folding assay (≈ 15 min^{-1}). Therefore, it is more likely that the nonlinearity of the Eyring plots is representative of a change in mechanism for folding rather than a change in the heat capacity.

In contrast to the other mutants, the temperature dependence of the P3 folding rate of the U167C mutant is similar to that of the wild type (fig. 17 B), with $\Delta H^{\ddagger}$ = 23.5±1.4 kcal/mol, yet the folding rate is ≅ 6 fold faster than in the wild type ribozyme. The $\Delta S^{\ddagger}$ = 12.7±4.8 cal/mol-K for the U167C mutant, and this increase in $\Delta S^{\ddagger}$ compared to the wild type value accounts for the rate acceleration.

A comparison of the temperature dependence of the wild type and mutant ribozymes supports the kinetic trap hypothesis for P3-P7 folding. The wild type ribozyme has a steep temperature dependence for P3 folding and a high activation enthalpy that represents the need to disrupt structure during the P3 folding. In contrast, the mutant ribozymes with accelerated folding have a decreased temperature dependence and activation enthalpy at temperatures above 22°C. However the mutations also seem to complicate P3 folding, folding in some mutants is biphasic and a fraction of the folding population folds slowly. A break in the temperature dependence at 22°C suggests a change in the folding mechanism at low temperatures. While the mutations do accelerate P3-P7 folding, they lead to a more complex kinetic folding pathway.

The effect of urea on P3-P7 folding.

Since mutations that destabilize native structure accelerate folding, it is logical that addition of denaturants, which also destabilize structure, to folding reactions would accelerate folding as well. The rate of P3-P7 folding in the wild type and mutant ribozymes was measured using the oligonucleotide hybridization assay in the presence of varying concentrations of urea. To ensure that the kinetic oligonucleotide hybridization assay accurately reflected folding in the presence of urea the effect of urea on probe binding was measured (fig. 12A). Urea does not appear to perturb probe binding or RNase H cleavage. The rate of P3 folding increases linearly with increasing urea concentration. In the presence 3 M urea, the rate of wild type P3 formation (k_{obs}=6 min^{-1}) is

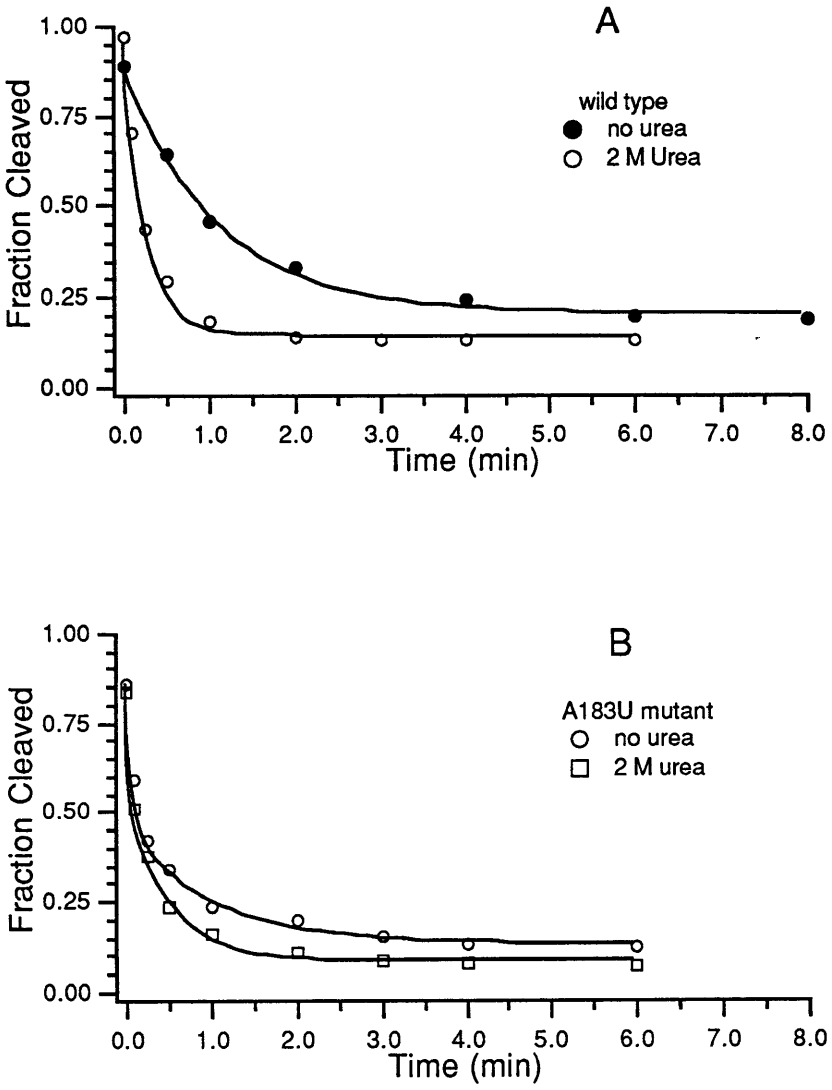


Figure 18. The effect of urea on P3 folding. A. Kinetic traces of wild type P3 folding in the presence of (●) no urea and (○) 2M urea. B. Kinetic traces of A183U mutant P3 folding in the presence of (○) no urea and (□) 2M urea.

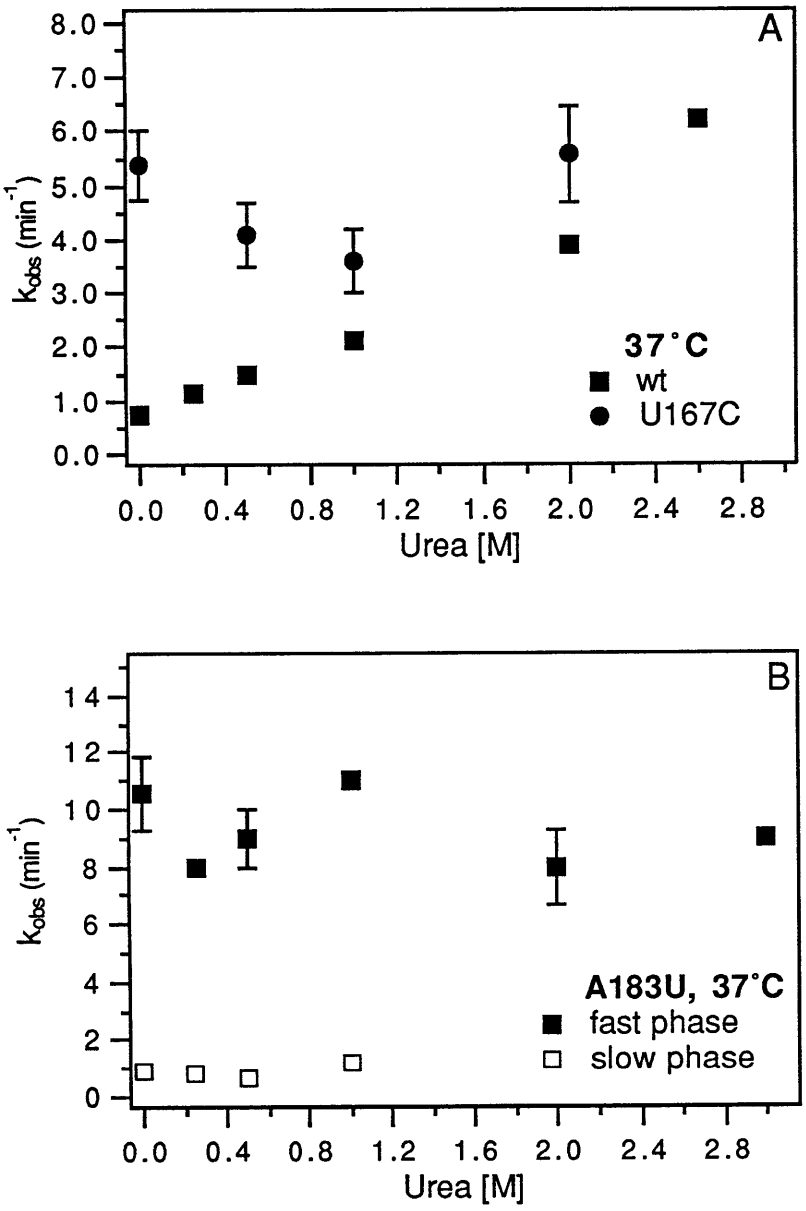


Figure 19. The effect of urea on the P3 folding rate. Error bars are less than or equal to the size of the marker unless otherwise shown. Each point represents at least two and as many as five experiments. Relative amplitudes for biphasic folding kinetics are given in table 2. (A) Wild type (■) and U167C (●), folding measured at 37°C. (B) A183U, folding measured at 37°C. Folding was biphasic at low concentrations of urea (■) fast phase rate, (❑) slow phase rate.

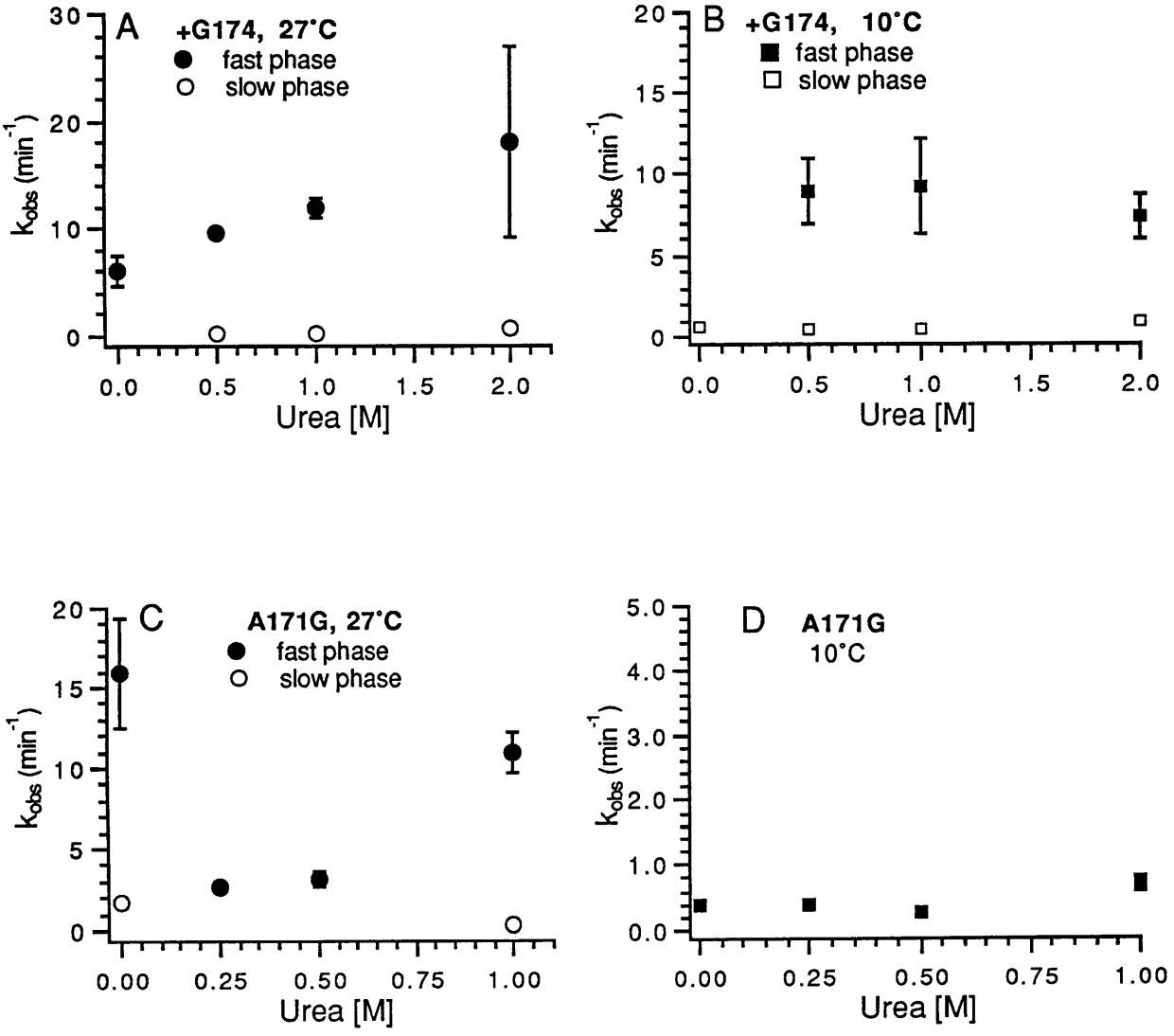


Figure 20: The effect of urea on P3 folding of (A) +G174, 27°C (B) +G174, 10°C (C) A171G, 27°C (D) A171G 10°C Where folding is biphasic the fast phases are represented by filled symbols, slow phases by open symbols.

Table 2. The fractional amplitude for folding in the presence of urea.

	A183U		+G174				A171G	
Urea [M]	37°C		27°C		10°C		27°C	
	fast phase	slow phase	fast phase	slow phase	fast phase	slow phase	fast phase	slow phase
0.0	.56±.03	.44±.03	1.0	0.0	0.0	1.0	.50±.06	.50±.06
0.25	.59±.05	.41±.04	n/a	n/a	n/a	n/a	1.0	0.0
0.5	.60±.04	.40±.04	.71±.04	.29±.2	.38±.04	.62±.02	1.0	0.0
1.0	.62±.12	.38±.08	.66±.06	.34±.08	.41±.06	.59±.04	.41±.07	.59±.01
2.0	1.0	0.0	.39±.06	.61±.04	.59±.3	.41±.2	n/a	n/a

The corresponding rate information for each mutant is shown in figures 19 and 20.

Table 3. The temperature dependence of the fractional amplitudes for P3 folding in 2M Urea.

Temp	wild type		A183U	
°C	fast phase	slow phase	fast phase	slow phase
45	1.0	0.0	n/a	n/a
37	1.0	0.0	1.0	0.0
32	1.0	0.0	0.49 ± 0.02	0.51 ± 0.02
27	0.43 ± 0.06	0.57 ± 0.06	0.47 ± 0.03	0.53 ± 0.04
22	0.56 ± 0.02	0.64 ± 0.02	0.38 ± 0.04	0.62 ± 0.1

The corresponding rate information for each ribozyme is shown in figure 21.

approximately 7 fold faster than in the absence of urea (k_{obs}=0.9 min^{-1}) (fig. 18A & 19A). Above 3M urea the ribozyme no longer folds to completion, although the P3 folding rate continues to increase. Thus, the effect of modest concentrations of denaturant is comparable to the effect of the fast-folding mutations.

The addition of urea should have little effect on the rate of P3 folding in the mutants because the folding intermediate is already destabilized. Urea does not appreciably affect the rate of P3 folding but it does affect the number of phases and the amplitudes of these phases of some mutants (fig. 18B, 19, 20). The P3 folding rate for U167C ($k_{obs} \approx 4$ min^{-1}) was approximately the same at all urea concentrations measured . At 37°C, the A183U mutant exhibits biphasic folding (fig. 18B & 19B), urea has little effect on the folding rate of either kinetic phase, but does affect the relative amplitudes of the two phases (table 2). The fast phase is favored slightly at urea concentrations between 0.25 and 2 M, and at concentrations $\geq$ 2 M folding occurs with a single fast phase. The presence of two folding phases indicates that there are two folding populations that fold with different rates. These populations may follow different folding pathways, or the slow folding population could be the result of a misfolded off-pathway intermediate that must slowly unfold before folding productively.

For the +G174 and A171G mutants, the effects of urea could not be measured at 37°C because the folding rates were already at our limit of detection ($\approx$15 min^{-1}). Instead, the folding rates were measured at two lower temperatures (10 and 27°C) that flank a break in the Eyring plot . In principle, if this break is due to a change in the folding mechanism, then urea may have different effects at the two temperatures. Figure 20 shows the effect of urea on the rate of P3 folding for the +G174 mutant, and table 2 gives the relative amplitudes of each phase. In the absence of urea, P3 folding in +G174 at both temperatures is best represented by a single phase. At 27°C folding is fast with a rate of $\cong$ 10 min^{-1} and at 10°C folding is slow with a rate of $\cong$ 0.5 min^{-1}. When urea is added folding becomes biphasic at both temperatures. The two folding populations fold with the

fast and slow rates measured in the absence of urea. At 27°C, increasing the concentration of urea increases the fraction of molecules folding with the slow rate. While at 10°C increasing the concentration of urea increases the fraction of molecules folding with the fast rate. Therefore it appears that, as with the A183U mutant ribozyme, there are two pathways by which P3 can fold in the +G174 mutant. This data supports the hypothesis that the break in the temperature dependence of the +G174 mutant ribozyme is caused by a change in the folding mechanism rather than a change in the heat capacity. The addition of urea allows molecules to fold through both pathways, and at 27°C the slow pathway is favored by addition of urea, while at 10°C the fast pathway is favored by the addition of urea.

The effect of urea on P3 folding of the A171G mutant is also complex (fig. 20). At 27°C, folding of P3 is biphasic, and urea affects the relative amplitudes of these phases. At 10°C, folding of P3 is monophasic and slow ($\approx$0.3 min^{-1}) at all concentrations of urea; however, only fifty percent of the molecules fold to completion. Thus, at high temperatures, the A171G mutant can fold through two pathways, and the fraction of molecules folding through each pathway is sensitive to urea. At low temperatures, however, P3 folding in the A171G mutant is unaffected by the addition of urea.

Urea and the mutations have similar effects on the energetic barriers to P3 folding.

To determine if urea has a similar effect on the energetic barriers to P3 folding as the selected mutations, the temperature dependence of wild type P3 folding was measured in the presence of 2 M urea and the activation enthalpy and entropy determined (fig. 21A and table 1). The activation enthalpy for folding is decreased by the addition of 2M urea, but the activation entropy also decreases. These changes are similar to those observed in the mutant ribozymes and further suggest that urea and the selected mutations act through the common mechanism of destabilizing an intermediate. At low temperatures ($\leq$ 27°C), the addition of 2M urea causes wild type P3 folding to become biphasic. As the temperature is decreased, the fraction of molecules folding through the slow pathway increases (table 3). The rate of the slow pathway for P3 folding

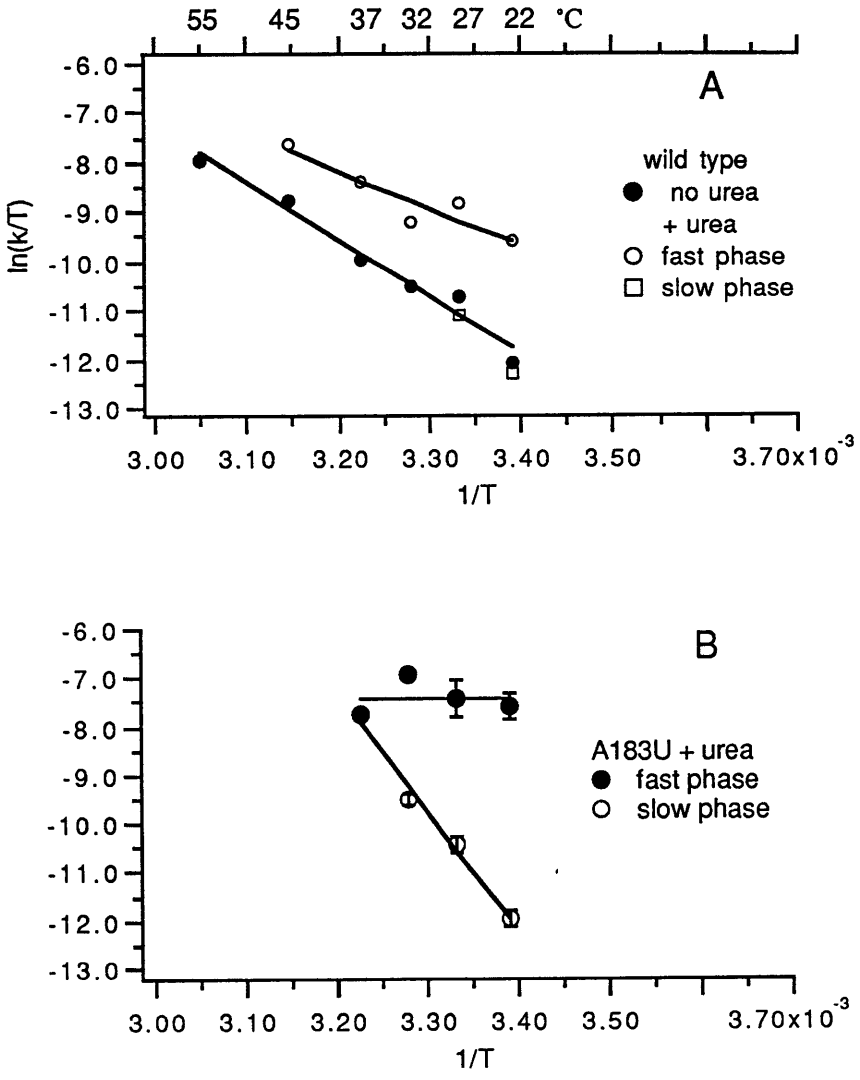


Figure 21. Eyring plot of the temperature dependence of P3 folding in the presence of urea. A. wild type B. the A183U mutant

at 27 and 22°C corresponds to the folding rate of the wild type ribozyme in the absence of urea. The biphasic folding curves indicate that urea reduces the kinetic barrier to P3 folding for only a fraction of the folding population at low temperatures.

Since urea and the mutations have similar effects on the energetic barriers to P3 folding, we wished to determine if urea further affected these barriers in the fast folding mutants. Therefore, the temperature dependence of A183U P3 folding in 2M urea was measured (fig. 21B). Below 37°C, the folding was biphasic in both the presence and absence of urea. The activation enthalpy of the fast phase of folding was unaffected by urea, but as the temperature decreased a greater proportion of the molecules folded through the slow pathway (table 3). The lack of change in the activation enthalpy for the fast folding phase indicates that the kinetic barrier to folding is unchanged by urea. However, the activation enthalpy of the slow phase of folding was increased by addition of urea (table 1). The increase in the activation enthalpy for the slow phase of P3 folding in the A183U mutant could indicate that urea destabilizes this transition state relative to the ground state.

The fast folding mutations reveal an additional slow step on the folding pathway.

In the proposed folding pathway for the *Tetrahymena* ribozyme, the slow rearrangement that leads to P3-P7 formation is the slowest step on the folding pathway. The selected mutations accelerate this slow rearrangement and we wished to determine if there is another slow step on the pathway or if overall folding is also accelerated by the mutations. To answer this question, a previously developed gain of activity assay was used to measure the rate of formation of the catalytically active structure (Zarrinkar & Williamson, 1994). In this assay, the fraction of active ribozyme present at various times after folding is initiated by Mg^{2+} addition is determined by the magnitude of the kinetic burst that corresponds to the first turnover in a multiple turnover cleavage reaction. The fraction of substrate cleaved as a function of folding time gives the rate of formation of the active structure.

All folding curves measuring the rate of gain of catalytic activity were best fit by a single exponential folding curve. Previous experiments found the rate of gain of activity for the wild type ribozyme to be 0.6 min^{-1}, similar to the rate measured for P3-P7 folding of 0.7 min-1. However, when these experiments were repeated and the folding time extended to thirty minutes the rate was measured to be 0.1 min^{-1}. When the rate of gain of activity was measured for mutant ribozymes with accelerated P3-P7 folding the rate was also measured to be 0.1 min^{-1} (fig. 22). This slow rate for the gain of catalytic activity strongly suggests the presence of an additional slow step on the folding pathway. However, the difference between the rate measured for gain of catalytic activity and the rate of P3-P7 folding in the wild type is still fairly small, and while the rate difference between the two steps in the mutant ribozymes is large the slow rate for gain of catalytic activity might be due to the presence of the mutations. Therefore, additional evidence was necessary to demonstrate that the two folding steps were distinct.

The presence of magnesium is required for folding of the *Tetrahymena* ribozyme. The folding pathway contains steps that are both magnesium dependent and magnesium independent. A difference in the magnesium dependence of the two slow steps would confirm that they were distinct. Therefore, the magnesium dependence of each step was measured and compared. The rate of formation of the P3-P7 domain in the wild type ribozyme is Mg^{2+}-independent (Zarrinkar & Williamson, 1994)(fig. 23A). Measurement of the Mg^{2+} dependence of folding to the catalytically active state revealed that the overall folding rate decreases as the concentration of Mg^{2+} is increased (fig. 23B). The difference in the Mg^{2+} dependence between P3 folding and folding to the catalytically active structure confirms the existence of a new folding step distinct from P3 folding.

To confirm that the new folding step is the same in the wild type and mutant ribozymes, the magnesium dependence of the rate of gain of activity was compared in the wild type and A183U ribozymes (fig. 23). Both the rates of folding to the active structure at all Mg^{2+} concentrations and the Mg^{2+} dependence of the rate was similar for the A183U mutant and wild type (fig. 23B).

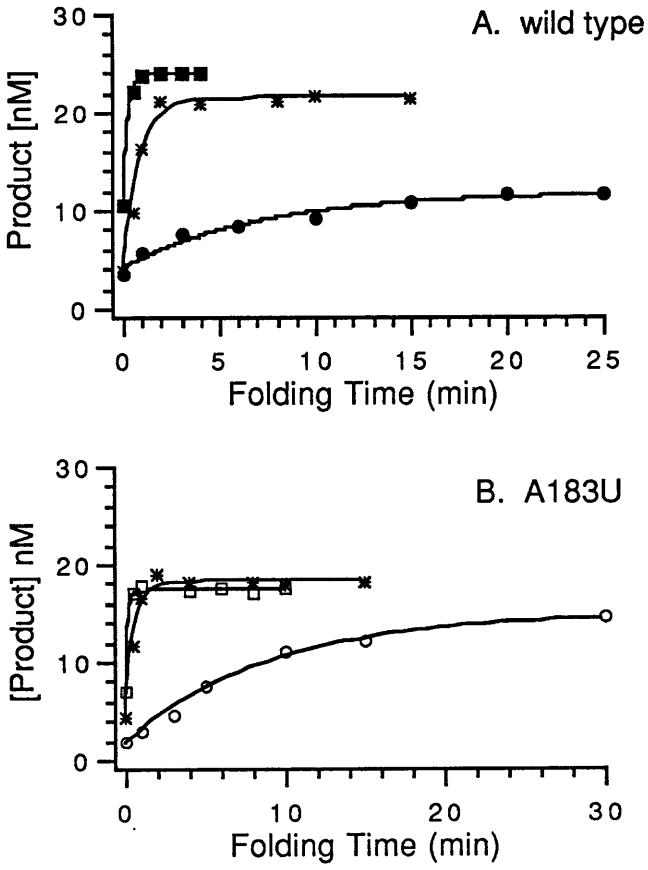


Figure 22: The rate of formation of the catalytically active structure as measured by monitoring the amount of substrate cleaved after folding in the presence of Mg^{2+} for the indicated periods of time. (A) Data for the wild type ribozyme. (●) folding in 10mM Mg^{2+}, (❑) folding in 10 mM Mg^{2+} in the presence of 2M urea, (✳) folding in 2mM Mg^{2+}. (B) Data for the A183U ribozyme. (●) folding in 10mM Mg^{2+}, (❑) folding in 10 mM Mg^{2+} in the presence of 2M urea, (✳) folding in 2mM Mg^{2+}.

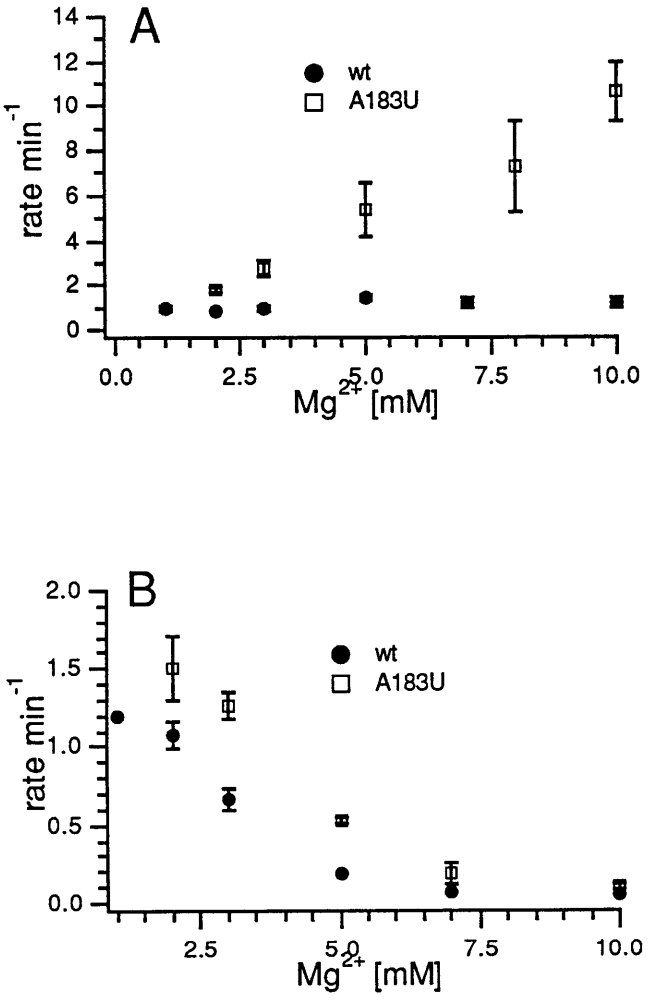


Figure 23: The Mg^{2+} dependence of the rate of P3 folding and folding to the catalytically active structure. A. The Mg^{2+} dependence of the rate of P3 folding. (●) wild type, (❑) A183U ribozyme. B. The Mg dependence of the rate of folding to the catalytically active structure. (●) wild type, (❑) A183U ribozyme

However, the rate of P3-P7 folding increases at higher Mg^{2+} concentrations in the A183U ribozyme, while the rate of P3-P7 folding in the wild type is Mg^{2+}-independent (Zarrinkar & Williamson, 1994). This suggests that the step measured by monitoring P3 folding is different in the two ribozymes, while the step measured by monitoring folding to the active structure is the same in the wild type and A183U mutant ribozyme. Therefore, the fast folding mutations do not fundamentally alter the folding pathway and they primarily affect the P3-P7 folding step.

The additional slow folding step is also the result of a kinetic trap.

Analysis of the fast folding mutants has demonstrated that the slow folding of the P3-P7 domain is the result of a kinetic trap. Is the slow folding rate of the final folding step also the result of a kinetic trap? A large temperature dependence can indicate the presence of a kinetic trap. Therefore, the temperature dependence of the folding rate constants for P3 folding and folding to the catalytically active structure were measured at 10 mM, where folding measured by gain of activity is slow, and 2 mM Mg^{2+}, where folding is fast. The data were analyzed using Eyring transition state theory to obtain values for the activation enthalpy and entropy for folding (fig 24 & table 4).

The activation enthalpy for both P3 folding and folding to the active structure in the wild type ribozyme at 10 mM Mg was high (fig. 24A and table 4), with $\Delta H^{\ddagger}$=23±0.6 kcal/mol for P3 folding, and $\Delta H^{\ddagger}$=18±6.0 kcal/mol for formation of the active structure. However, the activation entropy measured for each step was quite different (table 4), providing additional evidence that two different steps were monitored. When the Mg^{2+} concentration is lowered to 2 mM, the temperature dependence of the two steps becomes nearly identical with $\Delta H^{\ddagger}$=41±2.0 kcal/mol and $\Delta S^{\ddagger}$=66±7.0 cal/mol-K for P3 folding, and $\Delta H^{\ddagger}$=40±1.3 kcal/mol and $\Delta S^{\ddagger}$=63±4.1 cal/mol-K for formation of the active structure. The similar temperature dependence for the two measurements at 2 mM Mg^{2+} implies that at the lower Mg^{2+} concentration the rate of the final step is accelerated and that formation of P3-P7 becomes the slowest step on the folding pathway.

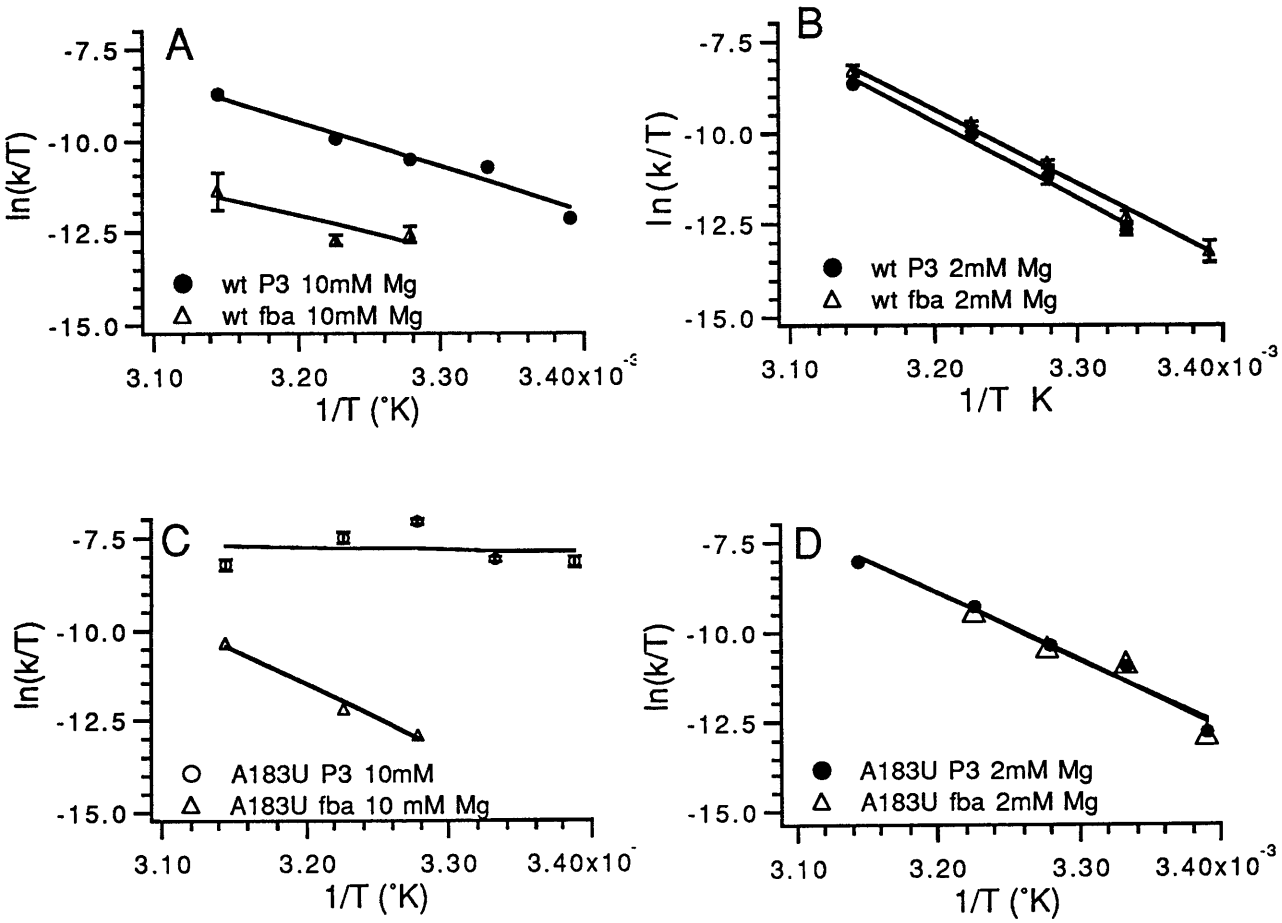


Figure 24: Temperature dependence of P3 folding and folding to the catalytically active structure at 10 mM and 2 mM Mg^{2+}. All points represent the average of at least three experiments, error bars are equal to or less than the size of the markers unless indicated. (A) wild type temperature dependence at 10 mM Mg^{2+} (●) P3 folding, (▲) folding to catalytically active structure (B) wild type temperature dependence at 2 mM Mg^{2+} (●) P3 folding, (▲) folding to catalytically active structure (C) A183U temperature dependence at 10 mM Mg^{2+} (●) P3 folding, (▲) folding to catalytically active structure (D) A183U temperature dependence at 2 mM Mg^{2+} (●) P3 folding, (▲) folding to catalytically active structure.

Table 4: The effect of magnesium on activation enthalpy and entropy .

	wild type		A183U	
$[Mg^{2+}]$	$\Delta H^{\ddagger}$ kcal/mol	$\Delta S^{\ddagger}$ cal/mol-K	$\Delta H^{\ddagger}$ kcal/mol	$\Delta S^{\ddagger}$ cal/mol-K
P3 10mM	23 ± 0.6	8.0± 1.9	-3.5± 1.1	-73 ± 3.5
fba 10mM	18 ± 6.0	-13 ± 12	38 ± 2.2	51 ± 7.0
P3 2mM	41 ± 2.0	66 ± 7.0	37 ± 2.6	53 ± 8.5
fba 2mM	40 ± 1.3	63 ±4.1	39 ± 5.0	60 ± 16

In the A183U mutant ribozyme at 10 mM Mg^{2+}, P3 folding and folding to the active structure have a very different temperature dependence. The activation enthalpy for P3 folding is -3.5±1.1kcal/mol and activation entropy is -73±3.5 cal/mol-K, while for formation of the active structure $\Delta H^{\ddagger}$=38±2.2 kcal/mol and $\Delta S^{\ddagger}$=51±7.0 cal/mol-K. This clearly demonstrates that two distinct folding steps are measured by monitoring P3 folding and folding to the active structure. In contrast, at 2 mM Mg^{2+} the temperature dependence of P3 folding and folding to the active structure are nearly identical with $\Delta H^{\ddagger}$=37±2.6 kcal/mol and $\Delta S^{\ddagger}$=53±8.5 cal/mol-K for P3 folding and $\Delta H^{\ddagger}$=39±5.0 kcal/mol $\Delta S^{\ddagger}$=60±16 cal/mol-K for folding to the active structure. Combined with the information that P3 folding in the A183U mutant ribozyme is also Mg^{2+} dependent, this data indicates that at high Mg^{2+} concentrations P3 folding is much faster than the step measured by gain of activity, but as the Mg^{2+} concentration is lowered P3 folding slows and gain of activity accelerates, and as a result P3 formation becomes the rate limiting step in folding to the active structure. It is also possible that at lower Mg^{2+} concentrations in the A183U mutant ribozyme, a step that occurs before P3 formation becomes rate limiting, however there is no clear evidence for any other slow steps before P3-P7 formation.

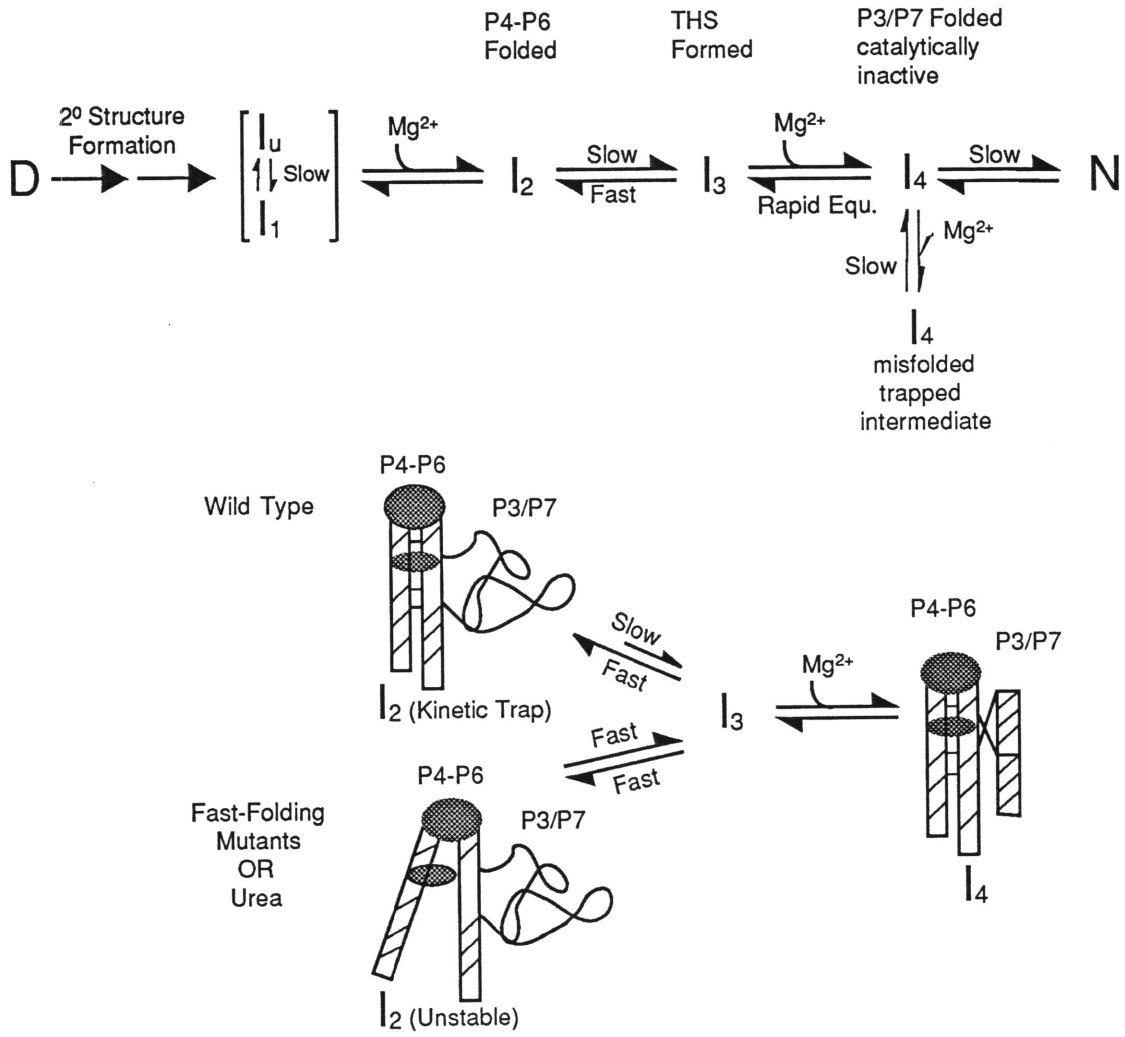


Figure 25. Modified kinetic folding pathway for the *Tetrahymena* ribozyme. The additional folding intermediate is labeled I_4. A model for the nature of the slow P3 folding step is shown under the folding pathway. Intermediate I_2 is stabilized by native contacts in P5abc, these contacts must be destabilized for productive P3-P7 formation.

The high activation enthalpy of the rate of gain of catalytic activity for the wild type and A183U mutant suggests that at 10 mM Mg^{2+} folding is slow due to the presence of a kinetic trap. This is supported by the acceleration of the rate of gain of catalytic activity in the presence of urea. This leads to a modification of the kinetic folding pathway shown in figure 25. There is a new intermediate on the folding pathway I_4, that represents an intermediate with folded P4-P6 and P3-P7 domains but that is not catalytically active. The decrease in the rate of gain of catalytic activity as the magnesium concentration increases suggests that this intermediate may contain misfolded structure that must be destabilized for the catalytically active structure to form.

The effect of Mg^{2+} on the folding pathway.

Most complex RNA tertiary folds require the presence of Mg^{2+}, which allows the close packing of the negatively charged phosphate backbone. In order to fold into it's complex, catalytically active structure the *Tetrahymena* ribozyme requires the presence of divalent metal ions, specifically Mg^{2+}. Mg^{2+} is believed to nucleate folding by allowing the formation of the P5abc subdomain (Cate et al., 1997), and additional Mg^{2+} ions are required to stabilize subsequent folding steps and are required for catalysis. These studies also show the importance of divalent metal ions, such as magnesium, to the stability of both non-native intermediates and the final folded structure. Magnesium plays a dual role in RNA folding, on the one hand Mg^{2+} is required for most complex RNA tertiary folds, but magnesium can also stabilize off-pathway intermediates that slow folding.

The work presented here highlights the importance of an appropriate Mg^{2+} concentration for folding. The concentration of Mg^{2+} used to initiate folding determines which of these steps is rate limiting, this is especially clear in the fast folding mutant A183U. In the wild type and A183U ribozymes, which folding step is rate limiting depends on the concentration of magnesium used to initiate folding. This leads to the observation that optimizing the rate of an RNA folding reaction may require a narrowly defined optimization of the Mg^{2+} concentration. The ability of RNA to form stable secondary structure interactions or small tertiary structures could potentially lead to the

presence of kinetic traps in RNA folding. It seems likely that Mg^{2+} is able to stabilize these traps as well as the native structure and can both accelerate and decelerate folding depending on the circumstances.

The *Tetrahymena* ribozyme has a rugged folding energy landscape.

Due to the complexity of the folding model, the kinetic folding pathway for the *Tetrahymena* ribozyme is best represented by a three-dimensional energy landscape (Dill & Chan, 1997). Such landscapes are based on a view of folding where molecules move down a funnel shaped energy landscape searching through possible configurations until the lowest energy state is reached. The vertical axis represents the free energy of the molecule and the radial axes represent the conformational coordinates. Hills or bumps in the landscape represent high energy structures, while valleys represent low energy structures or kinetic traps (Dill & Chan, 1997). A landscape is considered rugged if there are many such valleys and there are many intermediates populated, while it is considered smooth if kinetic traps are shallow and unpopulated. The utility of such landscapes lies in their ability to project two-dimensional linear folding pathways onto a three-dimensional surface, highlighting the existence of parallel folding pathways. The view of folding as an ensemble process does not conflict with instances where folding requires the sequential formation of particular structured intermediates. For example, in the *Tetrahymena* ribozyme a native P4-P6 domain is required to scaffold the formation of the P3-P7 domain (Doherty & Doudna, 1997; Zarrinkar & Williamson, 1996). While this intermediate must be formed for productive P3-P7 folding, it may not be populated if the formation of P3-P7 is fast.

A conceptualized energy landscape for the *Tetrahymena* ribozyme is shown in figure 26. In the wild type ribozyme, at least 90% of the folding molecules must go through an energy trap, I_2, which is represented by a red "moat" in the energy landscape. Studies on the folding of the *Tetrahymena* pre-RNA have shown that approximately 10% of the RNA population can fold rapidly within the thirty second dead time in a native polyacrylamide gel mobility assay

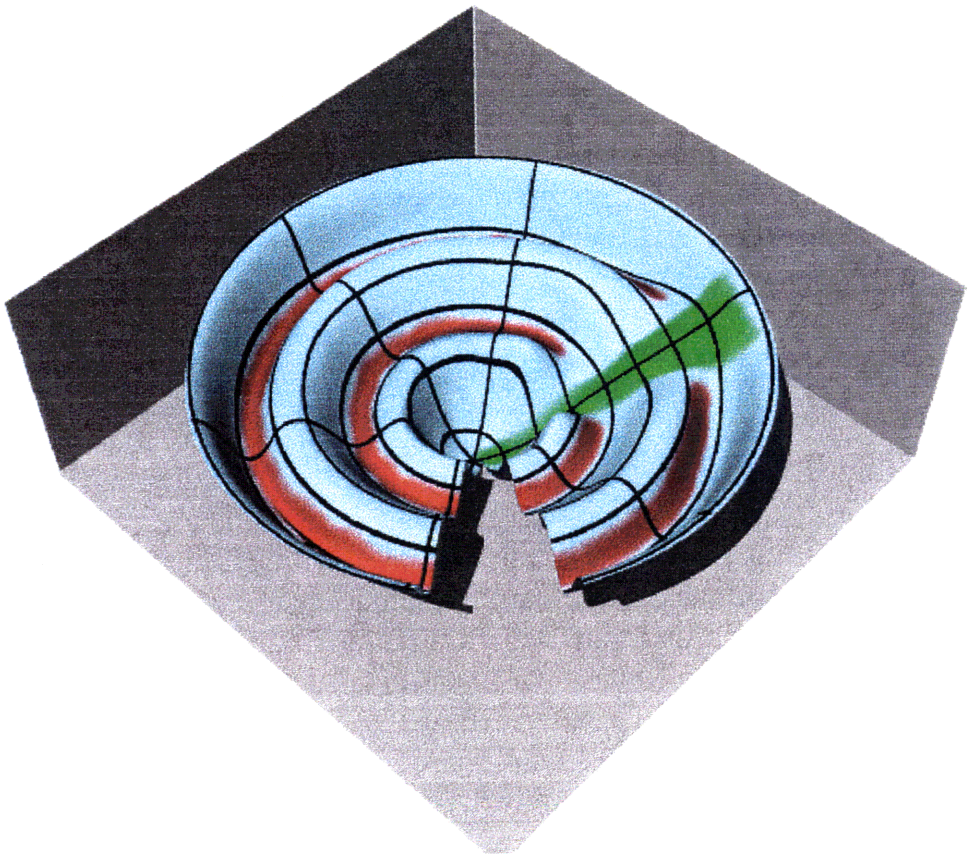


Figure 26. A conceptualized folding energy landscape for the *Tetrahymena* ribozyme. The green channel represents a fast pathway to the folded state, while the two red moats represent kinetic traps on the folding pathway.

(Pan et al., 1997). Therefore, a direct and rapid folding path to the native state may exist and is shown in green on the energy landscape. Destabilization of P4-P6 by mutation or denaturant either removes or reduces the depth of the moat and accelerates folding. However, under some conditions the moat is removed for only a subpopulation of molecules, such as for P3 folding in the wild type ribozyme at low temperatures in the presence of 2M urea. Under these conditions, P3 folding is biphasic, with a subpopulation still trapped by the low energy moat and another avoiding the trap. When an energy trap exists, it can obscure other folding steps that occur with an equal or faster rate. If, as is the case with the fast folding mutants, the energy trap is removed, an additional slow folding step may be revealed. The new slow step measured by the gain of activity assay is an example of this, and is represented as a second red moat on the landscape.

The existence of multiple kinetic traps shows that the *Tetrahymena* ribozyme has a rugged folding landscape. While mutations can reduce the depth of one folding moat, they can also allow additional kinetic traps to form and allow new pathways for folding to become populated, as is demonstrated by the biphasic P3 folding of some of the mutants. In contrast, denaturants such as urea usually accelerate folding, perhaps by globally reducing the depth of kinetic traps. Magnesium can have different effects on different regions of the energy landscape depending on how it effects the depth of kinetic traps or heights of activation barriers. In general, changes in folding conditions such as sequence changes, changes in temperature, or solvent conditions have dramatic effects on the folding energy landscape.

Kinetic traps are common in RNA folding.

The theory that folding of macromolecules occurs by the sequential formation of stable folding intermediates has dominated protein folding until recently, when it has been suggested that the presence of such stable intermediates may in fact slow the folding process (Fersht, 1995; Sosnick *et al.*, 1994). These stable intermediates are kinetic traps that must be destabilized for productive folding. It has long been known that tRNA folding from low salt conditions is slowed by the presence of misfolded secondary structure (Cole et al., 1972; Lynch & Schimmel, 1974).

However, it is only recently that studies have shown that the presence of intermediates stabilized by tertiary structure also slow folding, leading to the belief that kinetic traps may play an important role in RNA folding. An intermediate on the folding pathway of RNase P has been suggested to be the result of a kinetically trapped species (Pan & Sosnick, 1997). The folding of the *Tetrahymena* ribozyme pre-RNA has also been proposed to be slow due to the presence of a kinetically trapped intermediate (Pan et al., 1997). The results presented in this thesis demonstrate the presence of two kinetic traps on the folding pathway of the *Tetrahymena* ribozyme.

The presence of these kinetic traps leads to a refinement of the kinetic folding pathway, and a greater understanding of the barriers to folding for the two slow steps. In the wild type ribozyme after folding is initiated by the addition of magnesium, the P4/P6 region folds rapidly forming an intermediate I_2. The stable, rigid structure of P4-P6 restricts the conformational flexibility of I_2 and slows P3-P7 formation. When the mutations or urea destabilizes P4-P6, the kinetic trap is also destabilized allowing P3-P7 to form. The second slow step also has a high activation enthalpy and is accelerated by the presence of urea, two hallmarks of a kinetic trap. However, the observation that folding is slowed by increasing the concentration of magnesium suggests that this intermediate is a non-native, off pathway kinetic trap. These results lead to the view of RNA folding as the escape from a series of kinetic traps. Such a view explains the speed of folding of simple RNA structures such as tRNA and small ribozymes. More complex RNA structures, such as the *Tetrahymena* ribozyme and RNase P, are more likely to have a number of stable folding intermediates that act as kinetic traps and slow folding.

Similarities between RNA and protein folding.

It is interesting to compare and contrast the emerging view of RNA folding with the considerably more detailed view of protein folding. RNA folding occurs with much slower rates than protein folding. RNAs tend to have folding times on the order of minutes, while proteins fold on the second timescale, and a number of small proteins have been shown to fold in the millisecond time range. Such small, single domain proteins often display two-state kinetic folding

and appear to have smooth folding energy landscapes that mutations rarely complicate. One such protein is the chymotrypsin inhibitor 2 which folds with two-state kinetics and with a $t_{1/2}$ of 14 ms (Fersht, 1995). Small RNAs, like tRNA, can also fold on the millisecond timescale and may fold by similar mechanisms (Cole et al., 1972; Crothers et al., 1974). In larger, multi-domain proteins folding is more complex, and intermediates are often populated. The folding of larger RNAs, such as the *Tetrahymena* ribozyme, is much more complex than that of small proteins, and would be more analogous to larger multiple domain proteins. The folding pathway of lysozyme has striking similarities to that of the *Tetrahymena* ribozyme. Lysozyme is a 129 residue protein that consists of two structural domains, the α-domain and the β-domain, and folds by parallel slow and fast pathways. A folding intermediate with a folded α-domain slows the folding of the β-domain. It has been suggested that folding is slowed because the native-like α-domain restricts the conformational searching required for folding; however, the importance of non-native interactions in this intermediate is still a matter of debate (Miranker *et al.*, 1993; Radford *et al.*, 1992; Rothwarf & Scheraga, 1996; Wildegger & Keifhaber, 1997). This is analogous to a stable P4-P6 domain slowing the folding of the P3-P7 domain in the *Tetrahymena* ribozyme. It has recently been suggested that intermediates may in fact be kinetic traps that slow protein folding (Fersht, 1995; Sosnick et al., 1994), and it is beginning to appear that slow RNA folding is also due to the existence of kinetic traps (Pan et al., 1997; Pan & Sosnick, 1997; Treiber et al., 1997). Such traps may be the result of misfolded secondary structure (Cole et al., 1972; Woodson & Cech, 1991), or stable native-like folding intermediates that must be disrupted for productive folding (Treiber et al., 1997). The extremely slow rates for RNA folding compared with those of protein folding may be due to the greater stability of RNA secondary structure interactions and small tertiary structures, which makes it more difficult to escape from kinetic traps.

The finding that urea can accelerate folding by destabilizing a native-like folding intermediate, and that mutations that accelerate folding also destabilize this intermediate, suggests that populated intermediates are not necessary for productive RNA folding. This leads one to ask

why intermediates are found on the folding pathways of most large RNA molecules? It is likely that misfolded kinetic traps are common in large RNA molecules, however it is interesting that the *Tetrahymena* ribozyme folding pathway contains a native-like kinetically trapped intermediate. Point mutations can destabilize this intermediate and accelerate folding, but these mutations also lead to multiple folding phases that may represent parallel folding paths. All the mutations that accelerate folding display biphasic P3 folding kinetics. It is possible that the existence of such parallel paths is detrimental, and that funneling folding RNA molecules through a stable, native-like folding intermediate simplifies the folding landscape by decreasing the possibility of RNA molecules becoming trapped in stable misfolded structures. It will be interesting to determine if slow folding steps in other large RNA molecules are also caused by intermediates containing native-like folded structure.

3. Incorporation of a fluorescent probe into the *Tetrahymena* ribozyme.

There are relatively few effective methods for studying kinetic folding in RNA. Methods traditionally used in protein folding such as ultraviolet absorption and circular dichroism are of limited use in RNA. Ultraviolet spectroscopy in RNA is dominated by the hyperchromism due to secondary structure formation and while this can give information about secondary structure, little information about the more complex tertiary folding of RNA molecules can be gained. While circular dichroism has been used to study kinetic folding of RNase P (Pan & Sosnick, 1997), only information about global folding is gained and the structural nature of intermediates monitored is difficult to determine. Fluorescence spectroscopy is a powerful tool used in the study of protein folding, mainly because of the prevalence of tryptophan as a natural fluorophor. Unfortunately, standard RNA bases are not naturally fluorescent. Ribosomal RNA and tRNA contain modified bases, such as the Y-base, which are fluorescent however these bases are not common and are not necessarily located in regions of interest.

Fluorescence folding studies in tRNA.

The folding of tRNA has been more extensively studied than that of any other RNA. A number of these folding studies have utilized fluorescence spectroscopy using both fluorescent base analogs and external fluorescent probes. The fluorescent Y-base has been used to study the equilibrium magnesium binding and thermal unfolding of tRNA (Leroy *et al.*, 1977; Romer & Hach, 1975), and a fluorescent reporter molecule, 2-naphthoxy-acetic acid, has been attached to the amino group of isoleucyl-tRNA and used to monitor both equilibrium magnesium binding and magnesium induced kinetic folding (Lynch & Schimmel, 1974). This probe was found to be sensitive to magnesium induced tertiary folding, and based on the data a kinetic folding pathway was proposed (see chapter 1). The rare earth ion europium(III) has been used as a fluorescent

probe to monitor ion binding induced conformational changes (Wolfson & Kearns, 1975). These studies have shown that fluorescence spectroscopy can be a useful technique for studying folding and divalent metal binding in RNA.

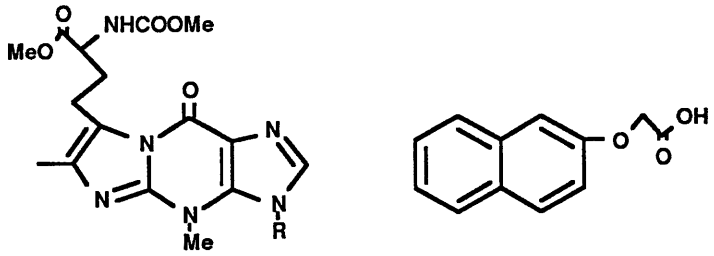


Fig. 27: A. The Y base. B. 2-naphthoxy-acetic acid

Site-specific incorporation of a fluorescent probe in the *Tetrahymena* ribozyme.

The *Tetrahymena* ribozyme is more than four times larger than tRNA and has a more complex tertiary fold. It is an ideal model for studying folding because this complex fold allows the formation of numerous folding motifs to be explored. However, the large size and complexity of the ribozyme also makes it more difficult to site-specifically incorporate a fluorescent probe. The ribozyme can be labeled at either the 5' or 3' ends in a manner similar to that used to label tRNA, but labeling the ends of the ribozyme does not give information about the folding of the catalytic core. The 3' end of the ribozyme is fairly unstructured making it difficult to assign kinetic events to particular structure formation. While the 5' end of the ribozyme has not been labeled, experiments studying the binding of fluorescently labeled substrate to the P1 helix and the subsequent docking of this helix onto the core of the ribozyme have been performed (Bevilacqua *et al.*, 1993; Bevilacqua et al., 1992). These studies have given valuable information about substrate binding, but is unlikely that 5' labeling of the ribozyme will yield any additional insights.

Therefore, we developed a system to allow introduction of fluorescent probes into specific folding domains so that the folding of these domains could be monitored.

To incorporate a fluorescent probe into the catalytic core of the *Tetrahymena* ribozyme, we expanded on a labeling system developed by Moore and Sharp, 1992. In this technique, modified RNA molecules were made by annealing transcripts primed with 2'-deoxy or methoxy substituted dinucleotides to a complementary DNA template and ligating the RNAs using T4 DNA ligase. In the modified system, three pieces of RNA are joined to form a full length RNA through two sequential ligation reactions. Two of the RNA strands are transcribed while the third, a small RNA containing the modified nucleotide, is chemically synthesized. The modified nucleotide first chosen was a commercially available deoxy-uridine with a five carbon linker that can be labeled with fluorescein (fig. 28 A).

The original ligation system was used on RNA having relatively little complex secondary structure. In contrast, the *Tetrahymena* ribozyme has a complex and stable secondary structure that

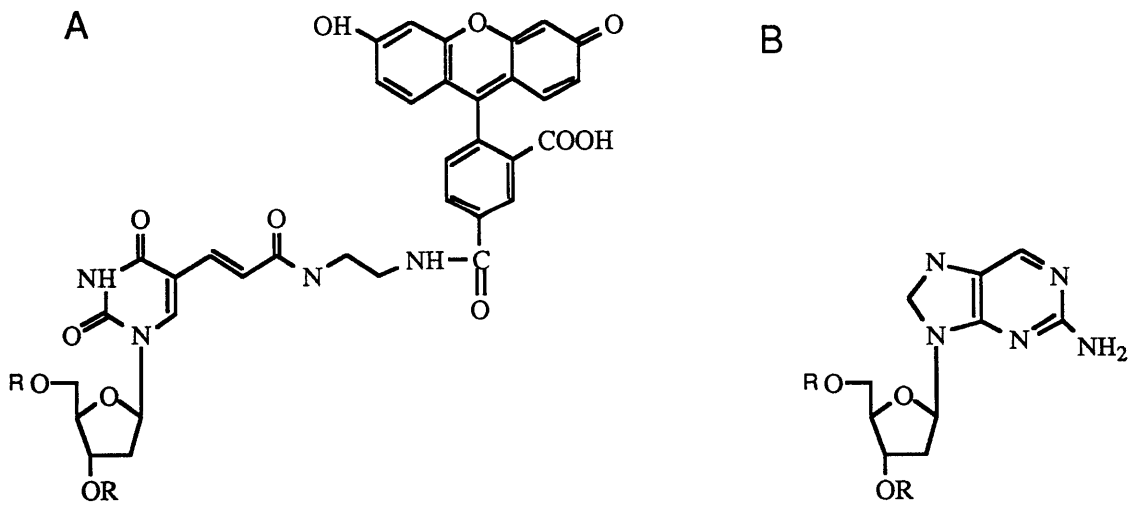


Figure 28: A. Structure of fluorescein labeled dU. B. Structure of deoxy-2-aminopurine.

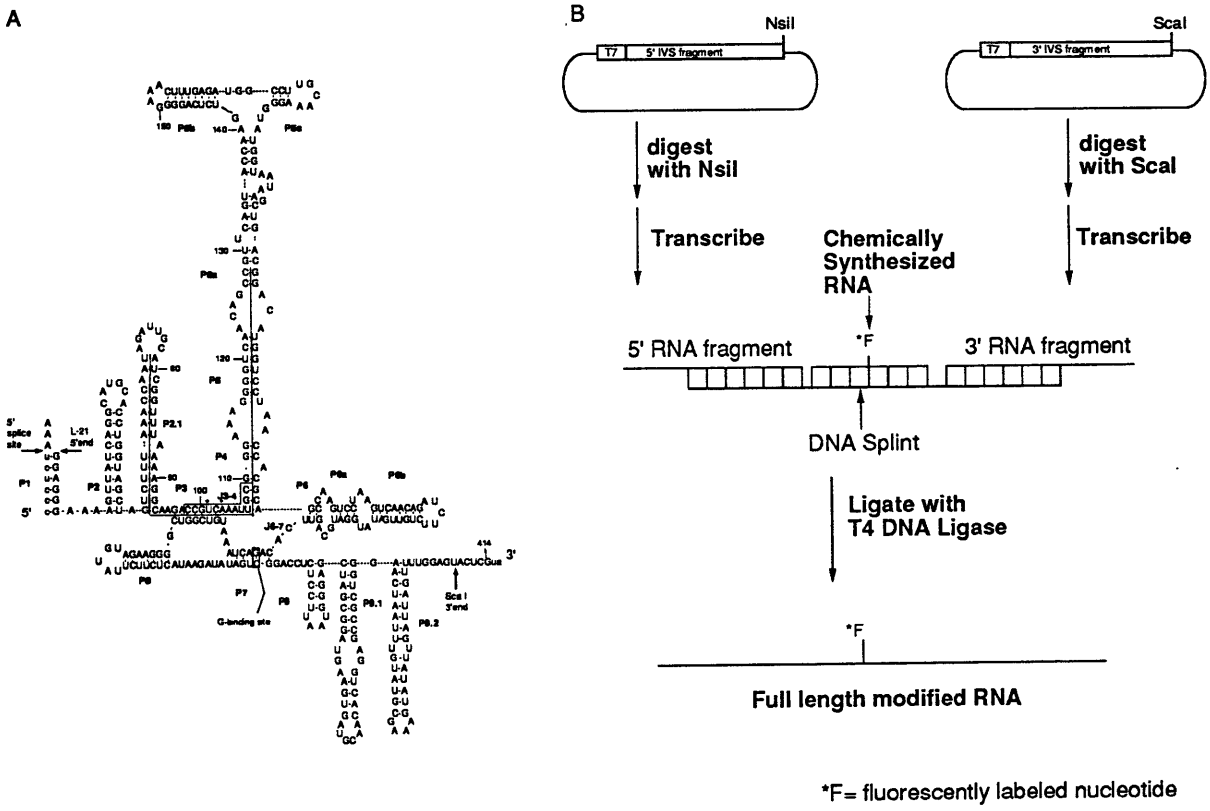


Figure 29: Ligation scheme for incorporation of a fluorescent probe into the core of the *Tetrahymena* ribozyme. A. The secondary structure of the ribozyme. The ligation splint is represented as a line, the sequence of the small synthetic RNA piece is boxed. B. Outline of the ligation procedure. Large RNA fragments are transcribed from plasmids L-21 Nsi I and P35' Sca I, transcripts and the labeled synthetic RNA are annealed to the DNA splint and ligated with T4 DNA ligase to give full length labeled ribozyme.

forms in the absence of divalent metals. Therefore, each of the ligation sites had to be chosen with care because stable secondary structures, such as short hairpins, would not anneal to the complementary DNA splint necessary to template the ligation. Additionally, the 3'-most large transcribed RNA fragment needed to start with two guanosines to ensure high transcription yields. Incorporation of a fluorescent probe into the P3 helix met these conditions, the P3 helix is a long rage interaction that does not form in the absence of magnesium, and a ligation junction could be placed in the middle of the 5' half of the P4 helix where a sequence containing three guanosines could be the start site for the transcription of the second transcribed RNA. Additionally, the folding of the P3 domain had been studied using the kinetic oligonucleotide hybridization assay allowing any results obtained from fluorescence experiments to be verified by a second folding assay. A forty-six nucleotide DNA splint complementary to the 3' half of P2, the 5' half of P3, and the 5' half of P4, P5, and P5a was synthesized (fig. 29). Plasmid templates for the transcription of the two RNA fragments were constructed. The template for the 5' fragment was cleaved by the Nsi I enzyme which led to the appropriate sequence for ligation at the 3' end. A twelve nucleotide synthetic RNA encompassing nucleotides 98-109 from the P3, J3/4, and P4 helices was chosen to contain the fluorescent probe. The modified base was substituted at position U101 to minimize any structural perturbation. U101 is a bulged base and any weakening of basepairing should not affect the normal ribozyme structure.

Trial ligations were performed using 5'-^{32}P labeled RNA fragments, ligated and unligated RNAs were separated on denaturing polyacrylamide gels and ligation yields were quantitated using phoporimager analysis. Even at the fairly unstructured P3 labeling site the ligation reaction was not efficient. The 3' ligation between the 3'-end of the small synthetic fragment and the 5'-end of P35' Sca I transcribed RNA was more efficient than the 5'-ligation between the 3'-end of L-21 Nsi I transcribed RNA and the 5'-end of the synthetic fragment. It is likely that the poor ligation yield was due to problems annealing the 3'-half of the P2 helix to the complimentary splint. The ligation yields were found to be highest when two sequential ligation reactions were performed. The two

ligation junctions were only twelve nucleotides apart and it may be that this did not leave sufficient room for two ligase molecules to bind the RNA-DNA helix. The 3'-ligation was performed first as it had a higher ligation yield of approximately thirty percent. The excess T4 DNA ligase was removed by phenol/chloroform extraction and the ligated RNA, RNA fragments, and DNA splint were purified by ethanol precipitation. The L-21 Nsi I RNA was added, the RNA fragments were annealed to the DNA splint and a second ligation was performed. The total yield averaged ten percent, and generated optical quantities of RNA.

Controls

The effect of magnesium on fluorescein.

Since we wished to use changes in the fluorescence intensity to monitor the magnesium dependent folding of the Tetrahymena ribozyme, it was important to determine what effect magnesium has on the fluorescence intensity of free fluorescein. A magnesium titration of free fluorescein is shown in figure 30A. As the magnesium concentration increases there is a decrease in the fluorescence intensity but this change is very rapid, in the dead time of the experiment. While magnesium did affect fluorescence intensity of free fluorescein, it did not appear to do so ona timescale relevant to folding. However, the fluorescein probe would be tethered to RNA not free in solution. Therefore, the effect of magnesium on the fluorescence of the small RNA fragment labeled both internally with the modified base and labeled at the 3' end was measured and is shown in figure 30 C and B respectively. Magnesium had no affect on the fluorescence of fluorescein conjugated with single stranded RNA, therefore, any changes in the fluorescence intensity of the full lenght labeled RNA should be due to RNA folding events.

The fluorescent probe does not perturb normal folding or catalytic activity.

Once a fluorescent probe was incorporated into the ribozyme, it was important to ensure that the probe did not disrupt the final folded structure or the folding process itself. The L-21 Sca I ribozyme used in these experiments catalyzes the cleavage of an eleven nucleotide RNA substrate.

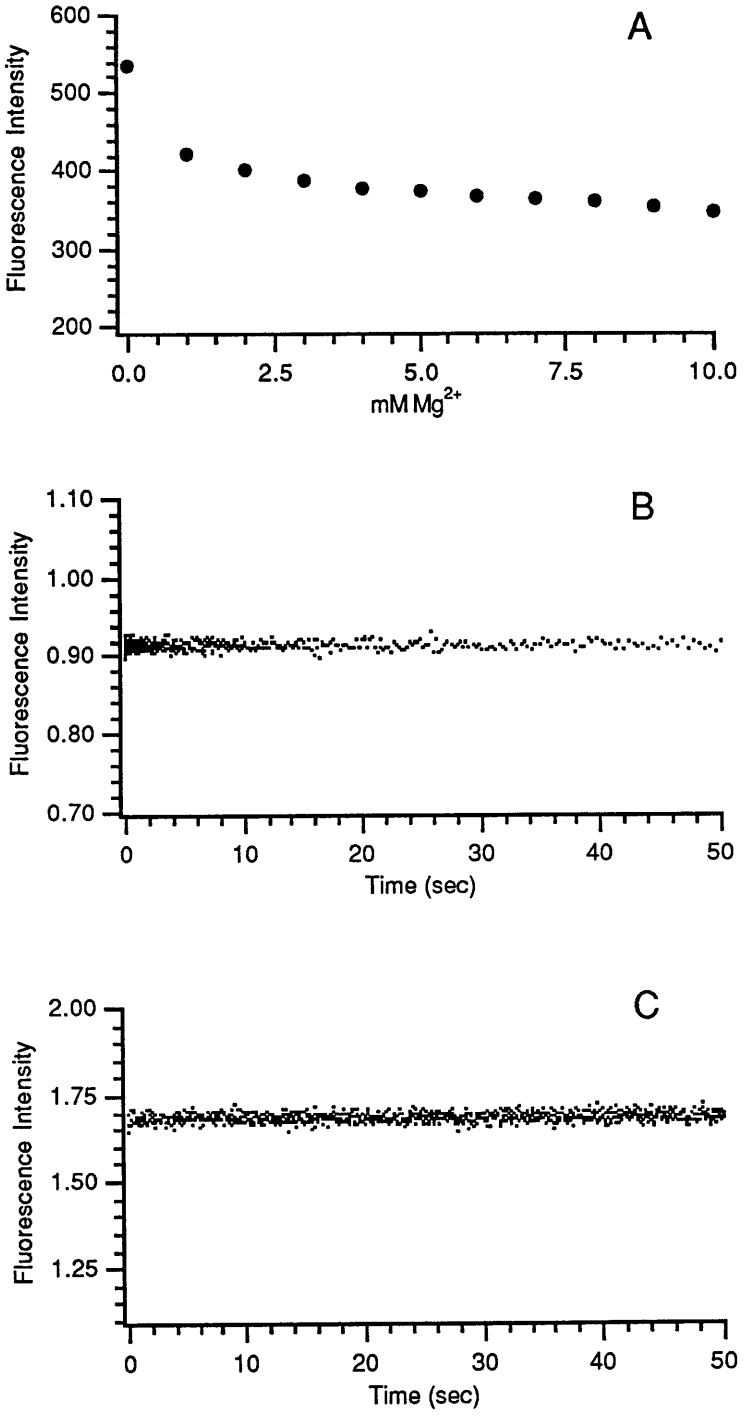


Figure 30. The effect of magnesium on fluorescein. A. Magnesium titration of free fluorescein. B. Kinetic trace of 3' fluorescein labeled synthetic RNA insert (CCGUCAAAUUGC-Fl) after addition of 10 mM Mg^{2+}. C. Kinetic trace of synthetic RNA insert labeled with a fluorescein modified nucleotide ($CCGU_{Fl}CAAAUUGC$) after addition of 10 mM Mg^{2+}.

This assay has been well characterized (Herschlag & Cech, 1990a), and was used to ensure that t

he final structure of the ribozyme was not perturbed. Under single turnover conditions, the rate limiting step in the cleavage reaction is substrate binding. The cleavage of ^{32}P labeled substrate under single turnover conditions of wild type ribozyme, and ribozyme labeled with fluorescein (fig. 31 A). Both k_{obsd} and k_{cat}/K_M were within error for the ribozymes demonstrating that the final structure of the ribozyme is not perturbed by the presence of the fluorescein probe.

The previously developed oligonucleotide hybridization assay was used to measure the rate of P3 folding for the labeled ribozyme. This assay uses the change in accessibility to complementary probes of the P3 helix after magnesium induced folding to obtain the folding rate. The rate for P3 folding measured for the fluorescein labeled ribozyme by the oligonucleotide assay was the same as for the wild type ribozyme (fig. 31 B).

The equilibrium magnesium binding of the *Tetrahymena* ribozyme has been studied using both hydroxyl-radical footprinting and oligonucleotide hybridization (Celander & Cech, 1991; Zarrinkar & Williamson, 1994). The magnesium binding monitored by fluorescein was compared with the binding monitored by oligonucleotide hybridization measuring P3 formation (fig. 32). The binding monitored by fluorescein labeled RNA had a similar magnesium midpoint as the wildtype ribozyme, but a lower Hill coefficient indicating that fewer magnesium ions are equired to saturate the fluorescence monitored. It may be that this probe can only monitor a subset of the binding events, or it may not accurately reflect P3 folding.

Kinetic folding monitored by fluorescence.

Fluorescein did not to perturb the folding pathway or the structure of the *Tetrahymena* ribozyme, but did it accurately monitor kinetic folding events? The folding of the fluorescein labeled ribozyme was studied using both steady state and stopped-flow spectroscopy. The RNA was heated at 95°C for 2 minutes and pre-equilibrated at 37°C for at least 3 minutes to correctly form the secondary structure and folding was initiated by the addition of magnesium. The steady

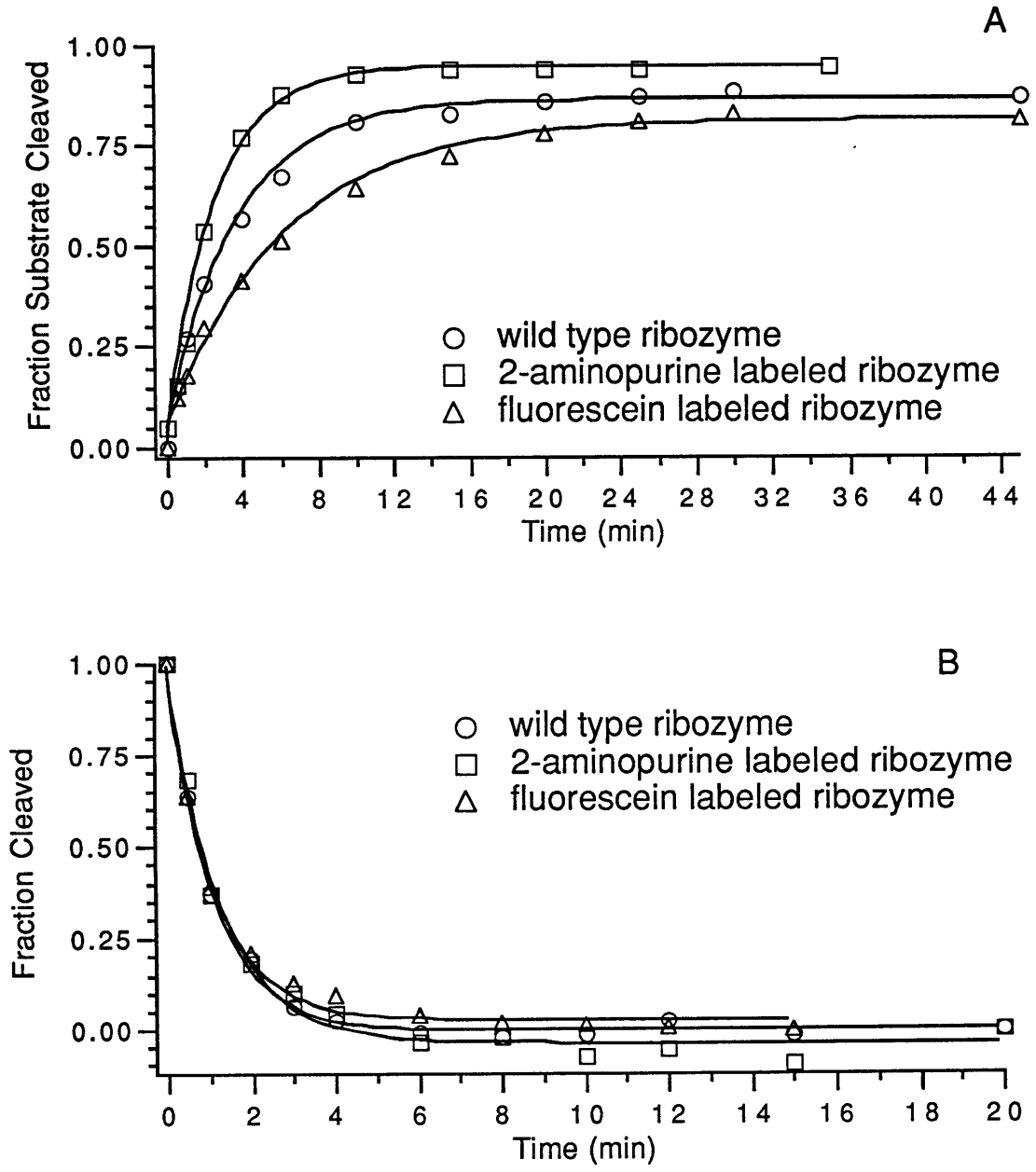
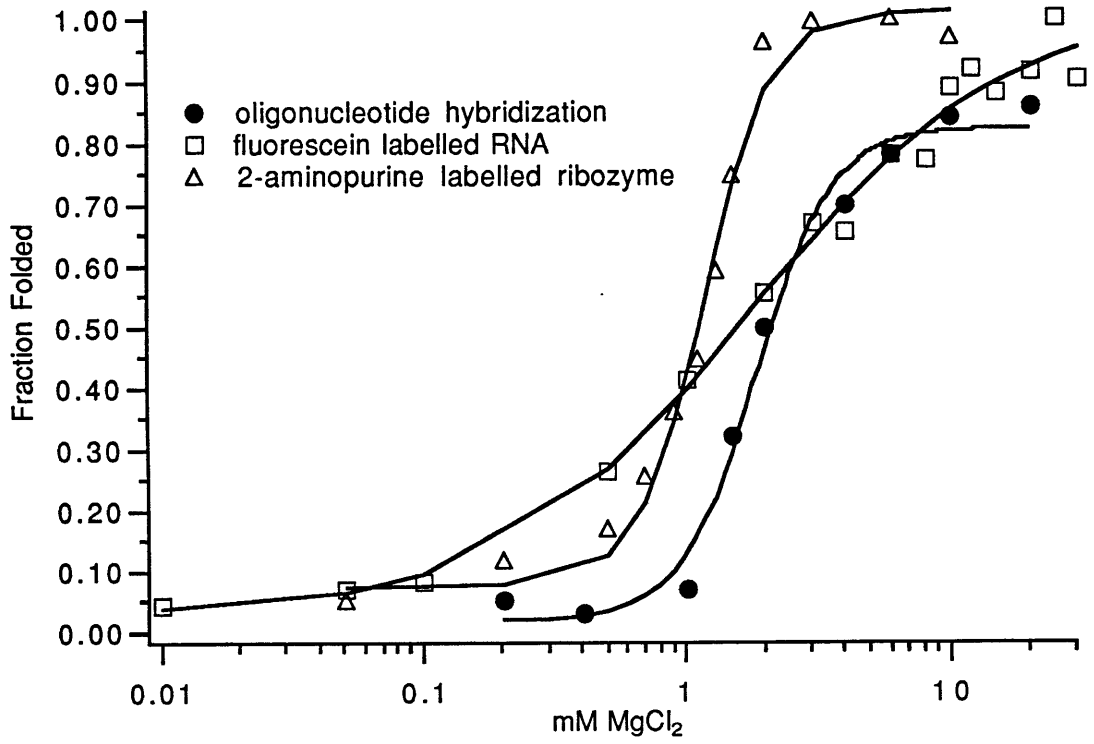


Figure 31: Control experiments to demonstrate that the fluorescent probes do not affect the final folded structure or the folding pathway. A. Single turnover activity assay. B. Oligonucleotide hybridization folding assay measuring the rate of P3 formation.



method	$Mg_{1/2}$	Hill Coefficient
oligonucleotide hybridization	1.8	3.0
fluorescien	1.7	0.90
2-aminopurine	1.2 mM	3.8

Figure 32: Equilibrium magnesium binding of the *Tetrahymena* ribozyme measured by oligonucleotide hybridization (●) fluorescence of the fluorescein probe (□), fluorescence of the 2-aminopurine probe (Δ). The midpoint for each magnesium titration and the Hill coefficient are listged below.

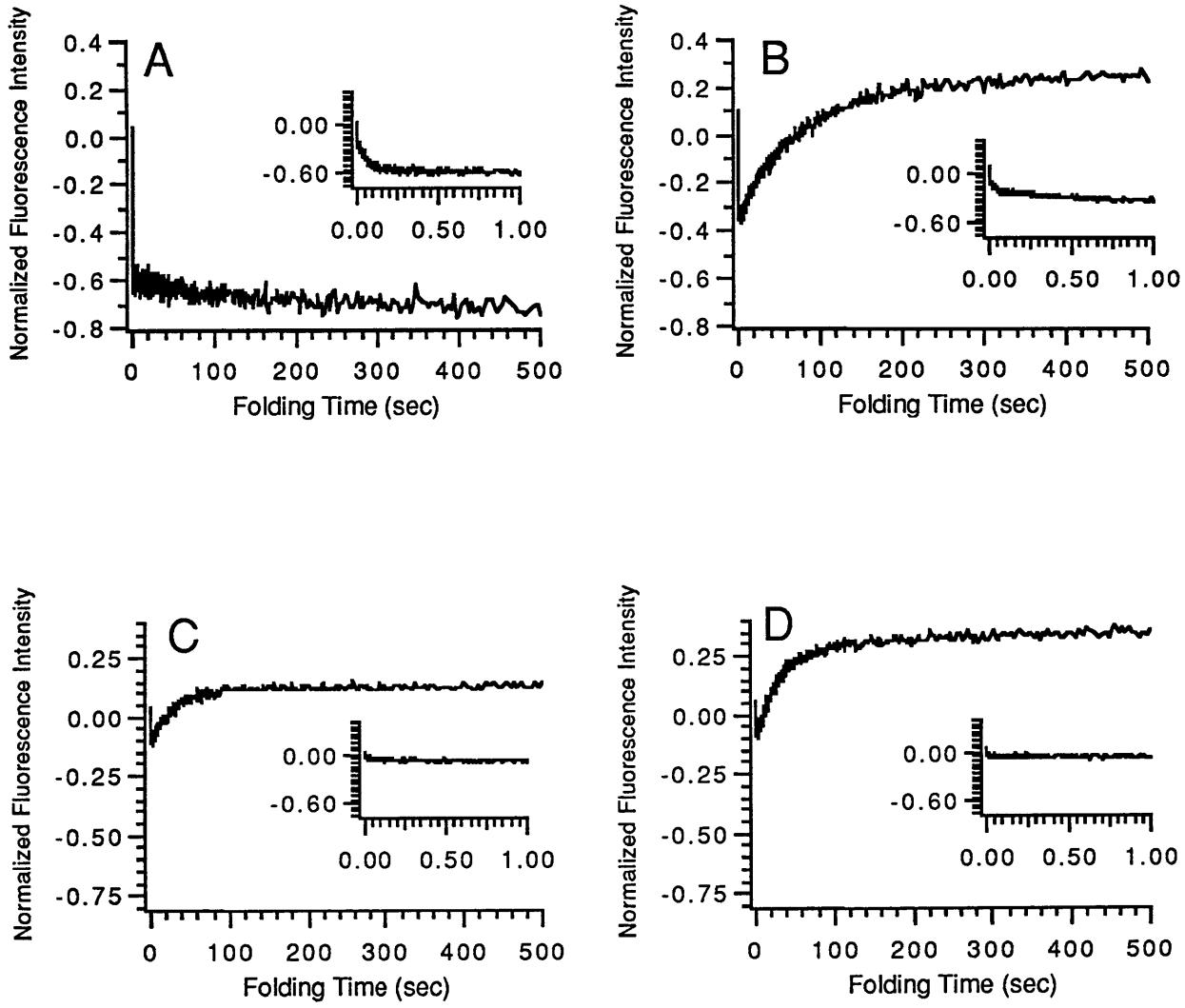


Figure 33: Stop flow fluorescence traces monitoring the magnesium induced folding of fluorescein labeled ribozyme. Inset shows the first second of the folding reaction. (A) 0.1 mM Mg^{2+} (B) 1 mM Mg^{2+} (C) 5 mM Mg^{2+} (D) 10 mM Mg^{2+}

state experiments had a dead time of approximately ten seconds. There was a burst in fluorescence intensity during this period followed by a slow folding phase with a rate of 1.75±.38 min^{-1}. This is very similar to the folding rate of .9 min^{-1} measured for P3 folding using the oligonucleotide hybridization assay. Stopped-flow spectroscopy was used to observe any folding events occurring in the dead time of the manual mixing experiments. When folding was initiated by the addition of 10 mM magnesium folding occurred in a single phase, and the burst in the dead time appeared not to correspond to any additional folding steps (fig. 33). However, when folding was initiated by lower concentrations of magnesium (≤ 1mM) a rapid decrease in the fluorescence intensity was seen on the second timescale. This fast phase is largest at 0.1 mM magnesium where the concentration is not sufficient for tertiary structure formation, and may simply reflect the effect of magnesium on the fluorescein probe seen with free fluorescein. While the similarity of the folding rates observed using fluorescence and oligonucleotide folding suggested that the two techniques were monitoring the same step, additional folding experiments using known folding mutations were performed. A mutation in the triple helical scaffold that aligns the P4-P6 domain with the P3-P7 domain is known to decrease the rate of P3-P7 folding to .1 min^{-1}. This mutant ribozyme was labeled with fluorescein and the folding rate was measured to be .9 min^{-1}. This rate combined with the differences in the equilibrium magnesium binding curves suggested that fluorescein was not monitoring P3 folding. However, it was not possible to determine what the change in fluorescence intensity reflected in terms of folding and magnesium binding. This result highlights the problem of assigning specific folding events with changes in fluorescence intensity. As we could not determine whether the fluorescence intensity changes were monitoring folding or an interaction of fluorescein with magnesium we decided not to pursue the use of fluorescein as a probe to monitor folding.

2-aminopurine as a probe to monitor folding.

Alternative probes with less chance of interactions with magnesium that may give false folding signals the fluorescent base were analyzed and the base analog 2-aminopurine was chosen

as a fluorescent probe (Figure 28B). While 2-aminopurine should be less likely than fluorescein to interact with magnesium, it has a much lower quantum yield and more RNA must be used to get as good a fluorescence signal. The fluorescence of 2-aminopurine is also quenched by stacking when it is incorporated in a double helix. Therefore we chose three positions in the P3 helix for incorporation of the probe to find the position which had the strongest signal. These positions were G100, U101 (a bulged nucleotide), and A102 (a nucleotide adjacent to the triple helical scaffold). The fluorescence intensity changes induced by the addition of 10 mM Mg^{2+} for both the full length ribozyme and the small synthetic RNA containing 2-aminopurine at each of these position is shown in figure 34. In the left hand column the fluorescence spectrum in the absence and presence of magnesium is shown for the full length ribozymes and in the right hand column the spectrum is shown for the synthetic RNA fragments. Magnesium increased the fluorescence of the three RNA fragments in the dead time of the experiments (10 seconds). The ribozyme with 2-aminopurine in the A102 position was the only ribozyme that showed time dependent changes in fluorescence intensity. It may be that being adjacent to the triple helical scaffold folding actually leads to the unstacking of this base which would account for the increase in fluorescence.

Before the 2-aminopurine labeled ribozyme could be used in folding assays controls needed to be performed to show that the probe did not affect the folding or final structure of the ribozyme. These were carried out in the same manner as with the fluorescein probe and are shown in figure 30. 2-aminopurine did not perturb the final structure or the folding rate measured by kinetic oligonucleotide hybridization. In addition the equilibrium magnesium binding was measured and closely resembled those measured using the oligonucleotide hybridization assay. Both the magnesium midpoint of the titration and the Hill coefficient for binding was very similar to the wild type, considerably better than that of the fluorescein probe.

Kinetic folding monitored by 2-aminopurine fluorescence.

The folding rate of the 2-aminopurine labeled ribozyme was measured at both 25 and 37°C (fig. 35). At 25°C, folding occurs in a single phase with a rate of .52±.03 min^{-1}. This is similar

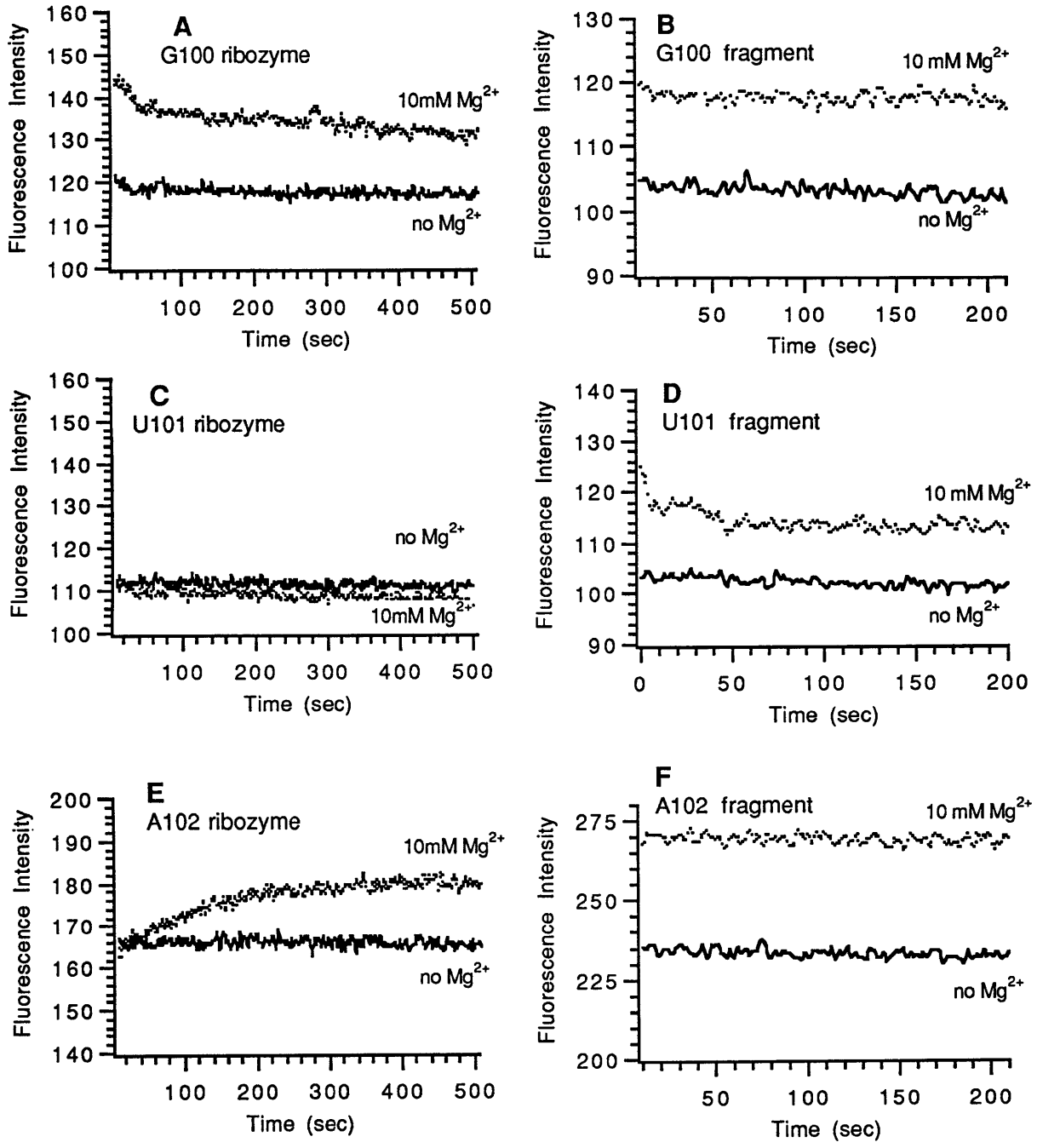


Figure 34. Kinetic traces of the fluorescence intensity change after the addition of 10 mM Mg2+. (A) 2-aminopurine at position G100 in the full lengthe ribozyme. (B) at position G100 in the synthetic RNA fragment. (C) at position U101 in the full length ribozyme. (D) at position U101 in the synthetic RNA fragment. (E) at postion A102 in the full length ribozyme (F) at postion A102 in the synthetic RNA fragment.

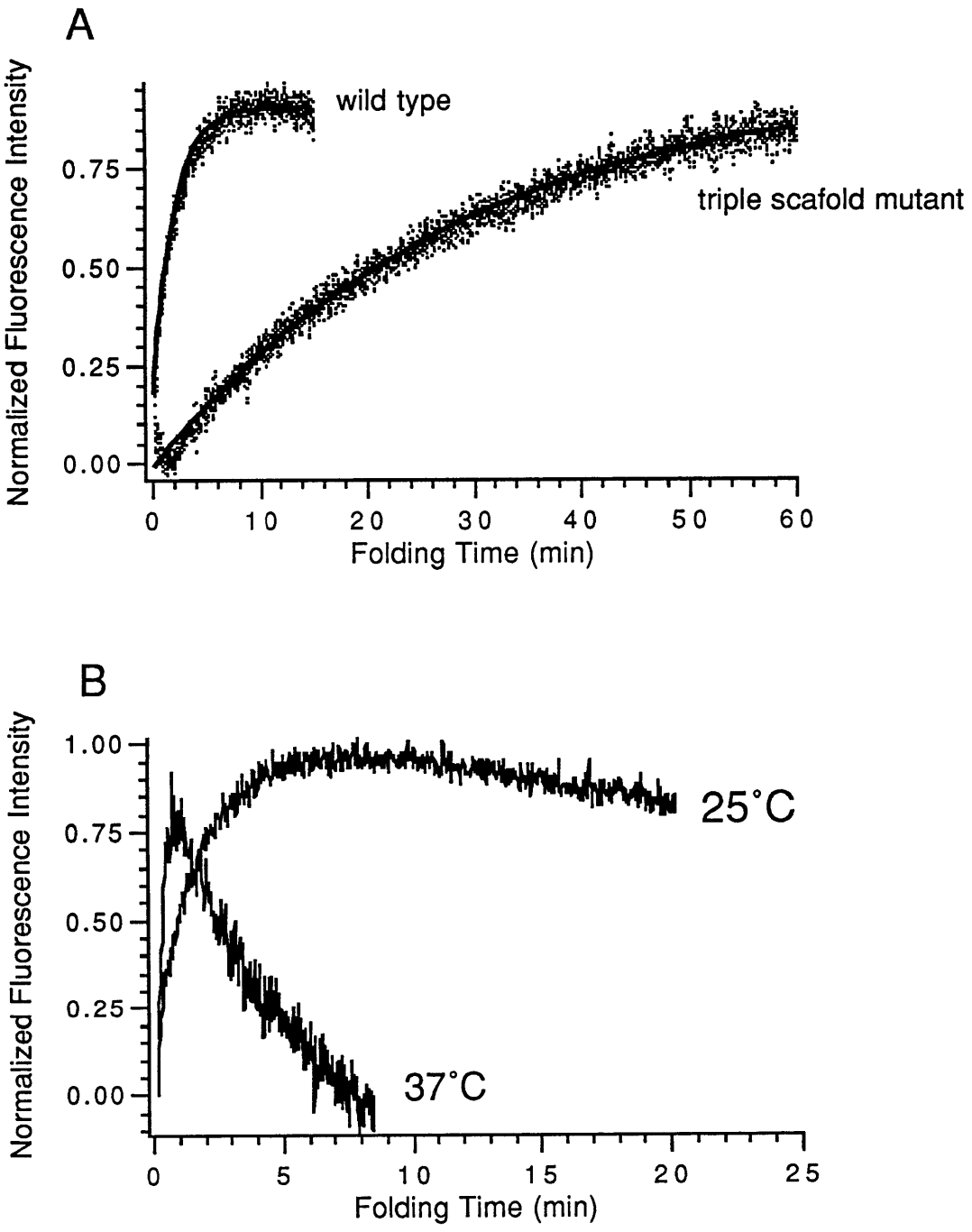


Figure 35: Fluorescence traces of magnesium induced folding of 2-aminopurine labeled ribozyme. A. Folding of wild type and triple helical scaffold mutant ribozymes at 25°C, 10mM Mg^{2+} B. Folding of wild type ribozyme at 37°C or 25°C, 10mM Mg^{2+}.

to the rate measured by oligonucleotide hybridization of .4±.05 min^{-1}. The folding of a ribozyme containing the mutation in the triple helical scaffold was also measured, and this mutant folded with a rate of .038±.003 min^{-1}. The decrease in the folding rate of the triple scaffold mutant ribozyme was good evidence that P3 folding was being monitored by the 2-aminopurine probe. At 37°C, a more complicated folding kinetics were measured. Folding was biphasic, with the fluorescence intensity first increasing with a rate of approximately 3 min^{-1} and then decreasing with a rate of approximately .1 min^{-1}. When folding at 25°C was monitored for a longer period of time a slow decrease in the fluorescence was observed as well. It appears that the 2-aminopurine probe monitors both P3 folding and an additional slow step. This may be due to 2-aminopurine being placed at A103 which is in the joining region between P3 and P4. While A103 is not a base that makes up the triple helical scaffold, it is directly adjacent to it. Therefore, folding monitored using fluorescence might be more sensitive to structural changes than the oligonucleotide hybridization assay.

Although the 2-aminopurine fluorescent probe appeared to monitor P3 folding, the existence of the additional slow phase once again highlighted the difficulty of assigning folding events to a particular fluorescence change. In addition the ligation efficiency for the 2-aminopurine probe was reduced compared to that with the fluorescein probe (2-5%) and as the quantum yield of 2-aminopurine is fairly low it was difficult to obtain optical quantities of labeled ribozyme. Therefore, we decided not to pursue fluorescence as a method of monitoring folding.

Summary

These experiments have demonstrated that modified nucleotides can be incorporated in the center of large complex RNAs. However, the number of sequences amenable to this technique are limited by secondary structures that may block annealing of the DNA splint necessary to template the ligation. In addition, low ligation yields can make it difficult to obtain large quantities of labeled RNA. This technique is best used on RNA molecules with less complex structure than the

Tetrahymena ribozyme where probes can be more easily incorporated at a number of positions in the RNA, or when only small quantities of RNA are required. The fluorescent base analog 2-aminopurine is a better reporter of folding than fluorescein. Perhaps the large size of fluorescein or its positioning, most likely in the minor groove due to the position of the linker off the C-5 of uridine, or the interaction of fluorescein with magnesium caused non-folding events to be monitored. However the low quantum yield of 2-aminopurine, and the quenching of fluorescence caused by base stacking limits its usefulness as a reporter of folding. Future experiments using fluorescence resonance energy transfer may be a better method of monitoring folding.

Appendix: Materials and Methods

Kinetic Oligonucleotide Hybridization Assay

RNA transcripts were prepared by transcription from plasmid pT7L-21, or fast folding mutants, linearized with Sca I. Reactions (50 µl) were performed for 2.5 hours at 37°C in 40 mM Tris-HCl (pH 7.5), 2 mM spermidine, 10 mM dithiothreitol, 25 mM MgCl2, 1 mM each GTP, CTP, UTP, 0.1 mM ATP and 50 µCi of [α-^{32}P] ATP (New England Nuclear) with 500 units of T7 RNA polymerase and 5 µg of template. The reaction was quenched with stop solution (90 mM EDTA and gel running dyes in 80 percent formamide) and the product purified on a six percent polyacrylamide gel and soaked in buffer containing 10 mM Tris-Cl (pH 7.5), 1 mM EDTA, and .3 M sodium acetate. After ethanol precipitation and resuspention in 1 X TE, RNA was quantitated by Cerenkov counting.

Folding was monitored after annealing ribozyme (final concentration 1 nM) in 60 µl of 1 X TE (pH 7.5) by heating to 95°C for 1 minute followed by incubation at 37°C for 3 minutes. When folding was monitored at a temperature other than 37°C, the ribozyme was first equilibrated at 37°C for 3 min. and then equilibrated at the folding temperature for an additional 3 min. Folding was initiated by addition of an equal volume of 2X folding buffer (1X: 50 mM Tris-HCl (pH 7.5), 10 mM NaCl, 1 mM dithiothreitol, $MgCl_2$ at the desired concentration). Aliquots (10 µl) were taken at the times indicated and added to 10 µl of 1X folding buffer containing oligonucleotide probe and RNase H (United States Biochemical, final concentration 0.1 U/µl) and enough $MgCl_2$ to bring the final concentration to 10 mM. Oligonucleotide binding and RNase H cleavage were allowed to proceed for 30 seconds before the reaction was quenched with 14 µl of stop solution. When the measurement of five, fifteen, and thirty second time points was necessary, each time point was measured in a separate reaction. Folding was initiated by adding 5 µl of annealed RNA to 5 µl of 2x folding buffer, and after the desired time 5 µl was removed and added to 5 µl of 1X

folding buffer containing oligonucleotide probe and RNase H. The zero time points were obtained by adding oligonucleotide probe and RNase H in 2X folding buffer and $MgCl_2$ (final concentration 10 mM) to the RNA immediately after annealing in a separate reaction. Products were separated on six percent denaturing polyacrylamide gels and quantitated using a Molecular Dynamics Phosphorimager. Folding was fit to single or double exponentials using Igor Pro to perform linear least squares analysis. The goodness of fit was determined from the Ψ^2 value calculated using the standard deviation from the average of at least three experiments as a weighting factor.

Kinetic Oligonucleotide Hybridization Assay in the Presence of Urea

Ribozyme (≈3 nM final concentration) was annealed in 24 µl of 1X Tris-EDTA by heating at 95°C for 1 min. and cooling at 37°C for 3 min. An equal volume of 2X folding buffer and urea to desired concentration was added. After folding for the desired time, 4 µl was added to 16 µl of quench (1X folding buffer containing 50 µM P3 oligonucleotide probe and 0.1U/µl RNase H) and after 30 seconds at 37°C, 14 µl of stop mix was added. Cleaved and uncleaved ribozyme were separated using a denaturing 6% polyacrylamide gel. The effect of urea on the Rnase H/oligonucleotide probe quench was determined by performing a urea titration and measuring the extent of cleavage at the zero time point. L-21 Sca I was annealed in 4 µl .1x tris-EDTA and between 0 and 5 M urea. The RNA was then added to 16 µl of quench solution allowed to react for 30 seconds and stopped by addition of stop mix. Cleavage was greater than 90% for urea concentrations up to 5 M urea (urea concentration in quench 1 M). At 5 M urea cleavage was still high (82%). Therefore, at all urea concentrations used in our assay the RNase H/oligonucleotide probe quench was unaffected by the presence of urea.

Measurement of formation of catalytically active ribozyme structure.

L-21 RNA was prepared for ribozyme activity assays by transcription from plasmid pT7 L-21, or fast folding mutants, linearized with Sca I. Reactions (100 µl) were performed for 3.5 hours at 37°C in 40 mM Tris-HCl (pH 7.5), 2 mM spermidine, 10 mM dithiothreitol, 25 mM $MgCl_2$, 4 mM each GTP, CTP, UTP, and ATP with 500 units of T7 RNA polymerase and 10 µg

of template. Formation of catalytcally active ribozyme structure was measured by annealing ribozyme (final concentration 20 nM) in 15 µl 10 mM Tris-HCl (pH 7.5), 2 mM NaCl, and adding an equal volume of 2X folding buffer at desired MgCl2 concentration. Folding was allowed proceed for the desired time before the endonuclease reaction was initiated by adding 15 µl of folding reaction to 15 ul of substrate mix containing 5'-^{32}P labeled substrate-CCCUCUAAAAA (final concentration 30 nM substrate, 500 mM GTP, 10 mM MgCl2, 10 mM NaCl). When assays were performed at temperatures other than 37°C the ribozyme was annealed by heating to 95°C for 1 min, incubated at 37°C for 3 min, and incubated at the folding temperature for 3 min, the endonuclease reaction was performed at 37°C. Products were separated on 20 percent denaturing polyacrylamide gels and the results were quantitated as above. The folding rate was determined by plotting the extent of substrate cleavage after 30 seconds as a function of the time the ribozyme was allowed to fold before substrate was added.

Construction of plasmids for three piece ligation system

Plasmids for the transcription of *Tetrahymena* ribozyme fragments for the three piece ligation system were constructed using standard PCR methodology. The plasmid encoding the 5'-fragment was constructed by performing PCR on the pT7L-21 plasmid using a 5'-primer complimentary to the T7 promoter sequence and the first 27 nucleotides of the L-21 Sca I sequence and containing a 5' overhang containing a Hind III restriction site, and a 3'-primer complimentary to nucleotides 73-99 and an overhang containing the Nsi I and BamHI restriction sites. PCR generated DNA was digested with Hind III and Bam HI, gel purified using agarose gel electrophoresis and purified using Promega Wizard PCR Preps, and ligated to PUC18 plasmid also digested with Hind III and Bam HI, gel purified, and dephosphorylated. The plasmid/PCR ligation mix was used in the chemical transformation of JM109 *E. Coli* and identified by sequencing performed using Sequenase kits (United States Biochemical/Amersham) and recommended protocols. The plasmid containing the 3' fragment was prepared in the same manner using a 5'-primer complimentary to nucleotides 110-135 and a 5' overhang containing T7

promotor sequence and a Bam HI restriction site, and a 3' primer complimentary to nucleotides 395-414 and an overhang containing a Hind III restriction site. Once individual transformants encoding the fragments were identified, a large quantity of the plasmids were prepared using Wizard Megaprep plasimd prep kits (Promega) and sequenced to ensure the proper plasmid was prepared.

Preparation of RNA for three piece ligation reaction

L-21 Nsi I RNA and P35' Sca I RNA were prepared by transcription from plasmids pT7L-21 Nsi I or pT7P35', linearized with Nsi I or Sca I respectively. Reactions (1 mL) were performed for 3.5 hours at 37°C in 40mM Tris-HCl (pH 7.5), 2 mM spermidine, 10 mM dithiothreitol, 25 mM $MgCl_2$, 4 mM each GTP, CTP, UTP, and ATP with T7 RNA polymerase prepared in the Williamson laboratory and 100 μg of template. The reaction was quenched with stop solution (90 mM EDTA and gel running dyes in 80 percent formamide) and the product purified on a six percent polyacrylamide gel and electro-eluted using an Elutrap apparatus. After ethanol precipitation and resuspention in 1 X TE, RNA was quantitated by measuring the OD_{260} and calculated extinction coefficients.

The fluorescently labeled RNA was prepared on a 1 μM scale using standard RNA synthesis techniques on an Applied Biosystems DNA synthesizer. Modified bases were included using deoxy-2-aminopurine phosphoramidite or a modified deoxy-UTP phosphoramidite containing a 5-carbon chain ending in a primary amine at C-5 (dU*) (both from Glenn Research). Synthesized RNA was phosphorylated using ATP and polynucleotide kinase (New England Biolabs). The dU* containing RNA was fluorescein labeled by incubating 1-10 nmols of RNA with approximately 100 nmols fluorescein isothiocyanate in 100 μl 100 mM NaCarbonate (pH 9), 75% dimethylformamide for 12 hours at 37°C in the dark. Excess fluorescein was removed by ethanol precipitation with five volumes of ethanol and labeled RNA was gel purified on a 20 percent denaturing polyacrylamide gel and eluted by soaking in 1X TE. Labeled RNA was desalted and further purified using C_{18} Sep-Pak Columns (Waters).

Ligation Reactions

Ligation reactions for the three RNAs were performed sequentially. P35'-Sca I RNA and the fluorescent insert RNA were annealed to the complimentary DNA splint in a 2:1:1 ratio by heating to 90°C for 2 minutes and slow cooling to 37°C over one hour. Ligase buffer (final concentration 66 mM Tris-HCl (pH 7.6), 6.6 mM MgCl2, 10 mM dithiothreitol, 66 μM ATP), 200 U T4 DNA ligase, and Rnasein ribonuclease inhibitor were added (final volume 300 μl) and the reaction was allowed to proceed at room temperature in the dark for 4.5 hours. The reaction was phenol/chloroform extracted to remove ligase and ethanol precipitated. L-21 Nsi I RNA was added to the first half ligation at a 2:1 molar ratio and annealed to the DNA splint by heating to 90°C for 2 minutes and slow cooling to 37°C over one hour. Ligation buffer and ligase were added as for the first ligation reaction and the reaction was allowed to proceed for 4.5 hours at room temperature in the dark. The reaction was phenol/chloroform extracted and purified on a 6 percent denaturing polyacrylamide gel. Full length, labeled RNA was isolated and soaked in 1XTE, .3 M sodium acetate and ethanol precipitated. The RNA was quantitated by OD_{260}.

Fluorescence Experiments

Fluorescence experiments were performed using either steady state or stopped flow fluorimetry. Steady state kinetic experiments had a dead time of approximately 10 seconds. Folding of fluorescently labeled ribozyme (10-100 nM, 500 μl) in 10 mM Tris-HCl (pH 7.5), 10 mM NaCl was initiated by the addition of $MgCl_2$ (10 μl) to a final concentration of 10 mM. Fluorescein fluorescence was monitored with an excitation wavelength of 495 nm and an emission wavelength of 520 nm. 2-aminopurine fluorescence was monitored with an excitation wavelength of 310 nm and emission wavelength of 370 nm. The fluorescence baseline in the absence of magnesium was subtracted from data and data were fit to either a single or double exponential as appropriate. Stopped flow experiments were only performed on the fluorescein labeled samples. Samples were mixed in a 1:1 ratio with a 20 mM $MgCl_2$, 20 mM Tris-HCl (pH 7.5), 20 mM NaCl

solution to give a final $MgCl_2$ concentration of 10 mM. Samples were excited at 480 nm and emission was monitored using a 495 nm bandpass filter.

Equilibrium magnesium binding was monitored by the serial addition of 1 µl of magnesium solutions at the appropriate concentration to give the desired final magnesium concentration to a solution of the labeled ribozyme (10-100 nM, 500 µl). Solutions were allowed to equilibrate until no more change occurred in the fluorescence spectra. The equilibrium endpoint from each magnesium concentration were fit to an expression of two state binding of n Mg2+ ions $[f=1/\{[Mg^{2+}]/[Mg^{2+}]_{1/2})^n+1\}]$, where f is the fraction fluorescence change $[(F_t-F_i)/(F_f-F_i)$ where F_t is the fluorescence intensity at time t, F_i is the initial fluorescence intensity, and F_f is the final fluorescence intensity], n is the number of magnesium ions bound per RNA molecule, and $[Mg^{2+}]_{1/2}$ is the midpoint of the transition.

References

Allain, F. H., Gubser, C. C., Howe, P. W., Nagai, K., Neuhaus, D. & Varani, G. (1996). Specificity of ribonucleoprotein interaction determined by RNA folding during complex formulation. *Nature* **380**, 646-650.

Altman, S., Kirsebom, L. & Talbot, S. (1995). Recent studies of RNase P. In *tRNA: Structure, Biosynthesis and Function* (Söll, D. & RajBhandary, U., eds.), pp. 67-78. ASM Press, Washington, D.C.

Baeyens, K. J., Bondt, H. L. D., Pardi, A. & Holbrook, S. R. (1996). A curved RNA helix incorporating an internal loop with G-A and A-A non-Watson-Crick base pairing. *Proc Natl Acad Sci* **93**, 12851-12855.

Banerjee, A. R., Jaeger, J. A. & Turner, D. H. (1993). Thermal unfolding of a group I ribozyme: the low-temperature transition is primarily disruption of tertiary structure. *Biochemistry* **32**, 153-163.

Barta, A., Steinier, G., Brosius, J., Noller, H. F. & Kuechler, E. (1984). Identification of a site on 23S ribosomal RNA located at the petidyl transferase center. *Proc. Natl. Acad. Sci.* **81**, 3607-3611.

Batey, R. T. & Williamson, J. R. (1996). Interaction of the *Bacillus sterothermophilus* ribozomal protein S15 with 16S RNA: II. Specificity determinants of RNA-protein recognition. *J Mol Biol* **261**, 550-567.

Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Nuhandiram, D. R., Kay, L. E., Frankel, A. D. & Williamson, J. R. (1996). A helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273**, 1547-1551.

Beaudry, A. A. & Joyce, G. F. (1992). Directed evolution of an RNA enzyme. *Science* **257**, 635-641.

Been, M. D. & Cech, T. R. (1986). One binding site determines sequence specificity of *Tetrahymena* pre-rRNA self-splicing, *trans*-splicing and RNA enzyme activity. *Cell* **47**, 207-216.

Berzal-Herranz, A., Joseph, S. & Burke, J. M. (1992). *In vitro* selection of active hairpin ribozymes by sequential RNA-catalysed cleavage and ligation reactions. *Genes Dev.* **6**, 129.

Bevilacqua, P. C., Johnson, K. A. & Turner, D. H. (1993). Cooperative and anticooperative binding to a ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8357-8361.

Bevilacqua, P. C., Kierzek, R., Johnson, K. A. & Turner, D. H. (1992). Dynamics of ribozyme binding of substrate revealed by fluorescence-detected stopped-flow methods. *Science* **258**, 1355-1358.

Bevilacqua, P. C. & Turner, D. H. (1991). Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2' OH groups. *Biochemistry* **30**, 10632-10640.

Blackburn, E. H. (1991). Structure and function of telomeres. *Nature* **350**, 569-573.

Blum, B., Sturm, N. R., Simpson, A. M. & Simpson, L. (1991). Chimeric gRNA-mRNA molecules with oligo(U) tails covalently linked at sites of RNA editing suggest that U addition occurs by transesterification. *Cell* **65**, 543.

Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak, J. W. & Tabak, H. F. (1987). Structural conventions for group I introns. *Nucleic Acids Res.* **15**, 7217-7221.

Caprara, M. G., Mohr, G. & Lambowitz, A. M. (1996). A tyrosyl-tRNA synthetase protein induces tertiary folding of the group I intron catalytic core. *J. Mol. Biol.* **257**, 512-531.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996a). Crystal Structure of a Group I Ribozyme Domain: Principles of RNA Packing. *Science* **274**(20), 1678-1695.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996b). RNA Tertiary Structure Mediation by Adenosine Platforms. *Science* **273**(20), 1696-1699.

Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nature Struct. Biol.* **4**(7), 553-558.

Cech, T. R. (1990). Self-splicing and enzymatic activity of an intervening sequence RNA from *Tetrahymena* (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **29**, 759-768.

Cech, T. R. (1993). Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. In *The RNA World* (Gesteland, R. F. & Atkins, J. F., eds.), pp. 239-269. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Cech, T. R., Zaug, A. J. & Grabowski, P. J. (1981). *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **27**, 487-496.

Celander, D. W. & Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science* **251**, 401-407.

Coetzee, T., Herschlag, D. & Belfort, M. (1994). *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 intron by acting as RNA chaperones. *Genes Dev.* **8**, 1575-1588.

Cole, P. E. & Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid: Relaxation kinetics of the early melting transition of methionine transfer ribonucleic acid. *Biochemistry* **11**, 4368-4374.

Cole, P. E., Yang, S. K. & Crothers, D. M. (1972). Conformational changes of transfer ribonucleoc acid: Equilibrium phase diagrams. *Biochemistry* **11**, 4358-4368.

Couture, S., Ellington, A. D., Gerber, A. S., Cherry, J. M., Doudna, J. A., Green, R., Hanna, M., Pace, U., Rajagopal, J. & Szostak, J. W. (1990). Mutational analysis of conserved nucleotides in a self-splicing group I intron. *J. Mol. Biol.* **215**, 345-358.

Crick, F. H. C. (1968). The origin of the genetic code. *J. Mol. Biol.* **38**, 367-379.

Crothers, D. M. & Cole, P. E. (1978). Conformational changes of tRNA. In *Transfer RNA* (Altman, S., ed.), pp. 196. MIT Press, Cambridge.

Crothers, D. M., Cole, P. E., Hilbers, C. W. & Shulman, R. G. (1974). The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. *J. Mol. Biol.* **87**, 63-88.

Dabbs, E. R. (1991). Mutants lacking individual ribosomal proteins as tools to investigate ribosomal properties. *Biochimie* **73**, 639-645.

Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzochcio, C. (1982). Making ends meet: A model for RNA splicing in fungal mitochondria. *Nature* **300**, 719-724.

Dieckmann, T., Suzuki, E., Nakamura, G. K. & Feigon, J. (1996). Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* **2**, 628-640.

Dill, K. A. & Chan, H. S. (1997). From Levinthal pathways to funnels. *Nature Struct. Biol.* **4**(1), 10-19.

Doherty, E. A. & Doudna, J. A. (1997). The P4-P6 domain directs higher order folding of the *Tetrahymena* ribozyme core. *Biochemistry* **36**, 3159-3169.

Doudna, J. A. & Cech, T. R. (1995). Self-assembly of a group I intron active site from its component tertiary structural domains. *RNA* **1**, 36-45.

Doudna, J. A. & Szostak, J. W. (1989). Miniribozymes, small derivatives of the *sunY* intron, are catalytically active. *Mol. Cell. Biol.* **9**, 5480-5483.

Downs, W. D. & Cech, T. R. (1990). An ultraviolet-inducible adenosine-adenosine cross-link reflects the catalytic structure of the *Tetrahymena* ribozyme. *Biochemistry* **29**, 5605-5613.

Downs, W. D. & Cech, T. R. (1996). Kinetic pathway for folding of the *Tetrahymena* ribozyme revealed by three UV-inducible crosslinks. *RNA* **2**, 718-732.

Ekland, E. H. & Bartel, D. P. (1996). RNA-catalyzed RNA polymerization using nucleoside triphosphates. *Nature* **382**, 373-376.

Emerick, V. L. & Woodson, S. A. (1993). Self-splicing of the *Tetrahymena* pre-rRNA is decreased by misfolding during transcription. *Biochemistry* **32**, 14062-14067.

Emerick, V. L. & Woodson, S. A. (1994). Fingerprinting the folding of a group I precursor RNA. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9675-9679.

Famulok, M. (1994). Molecular recognition of amino acids by RNA-aptamers: an L-citrulline binding RNA motif and its evolution into an L-Arginie binder. *J Am Chem Soc* **116**, 1698-1706.

Fersht, A. R. (1995). Optimization of rates of protein folding: The nucleation condensation mechanism and its implications. *PNAS* **92**, 10869-10873.

Fountain, M. A., Serra, M. J., Krugh, T. R. & Turner, D. H. (1996). Structural features of a six nucleotide RNA hairpin loop found in ribosomal RNA. *Biochemistry* **35**, 6539-6548.

Franceschi, F. J. & Nierhause, K. H. (1990). Ribosomal proteins L15 and L16 are mere late assembly proteins of the large ribosomal subunit. *J. Biol Chem* **265**, 16676-16682.

Frankel, A. D. (1992). Activation of HIV transcription by Tat. *Curr. Opin. Gen. Dev.* **2**, 293-298.

Giegé, R., Puglisi, J. D. & Florentz, C. (1993). tRNA structure and aminoacylation efficiency. *Prog. Nuc. Acids Res. Mol. Biol.* **45**, 129-206.

Gralla, J. & Crothers, D. M. (1973). Free energy of imperfect nucleic acid helices. II. Small hairpin loops. *J. Mol. Biol.* **73**, 497-511.

Green, R. & Szostak, J. W. (1994). *In vitro* genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the *sunY* group I self-splicing intron. *J. Mol. Biol.* **235**, 140-155.

Guerrier-Takada, C. & Altman, S. (1984). Catalytic activity of an RNA molecule prepared by transcription *in vitro*. *Science* **223**, 285-286.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857.

Guo, Q. & Lambowitz, A. M. (1992). A tyrosyl-tRNA synthetase binds binds specifically to the group I intron catalytic core. *Genes Dev.* **6**, 1357-1372.

Guthrie, C. (1991). Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* **253**, 157.

Harris, M. E., Nolan, J. M., Malhatra, A., Brown, J. W., Harvey, S. C. & Pace, N. R. (1994). Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J.* **13**, 3953-3963.

Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**, 20871-20874.

Herschlag, D. & Cech, T. R. (1990a). Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry* **29**, 10159-10171.

Herschlag, D. & Cech, T. R. (1990b). Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 2. Kinetic description of the reaction of an RNA substrate that forms a mismatch at the active site. *Biochemistry* **29**, 10172-10180.

Herschlag, D., Khosla, M., Tsuchihashi, Z. & Karpel, R. L. (1994). An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J.* **13**, 2913-2924.

Jacob, W. F., Santer, M. & Dahlberg, A. E. (1987). A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins. *Proc. Natl. Acad. Sci.* **84**, 4762-4766.

Jacquier, A. (1997). RNA-catalyzed RNA processing. In *Eukaryotic mRNA processing*. (Krainer, A. R., ed.), pp. 1-28. IRL Press, Cold Spring Harbor.

Jenison, R. D., Gill, S. C., Pardi, A. & Polinsky, B. (1994). High resolution molecular discrimination by RNA. *Science* **263**, 1425-1479.

Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H. J., Seeman, N. C. & A.Rich. (1974). Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* **185**, 435-440.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. & Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147-157.

Laird-Offringa, I. & Belasco, J. G. (1995). Analysis of RNA-binding proteins by *in vitro* genetic selection: Identification of an amino acid residue important for locking U1A onto it's RNA target. *Proc. Natl. Acad. Sci.* **92**, 11859-11863.

Lambowitz, A. M. & Pearlman, P. S. (1990). Involvement of aminoacyl-tRNA sythetases and other proteins in group I and group II intron splicing. *TIBS* **15**, 440-444.

Latham, J. A. & Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science* **245**, 276-282.

Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (1996). New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID; a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chemistry and Biology* **3**, 993-1009.

Leroy, J.-L., Gueron, M., Thomas, G. & Favre, A. (1977). Role of divalent ions in folding of tRNA. *Eur. J. Biochem.* **74**, 567-574.

Limmer, S., Reil, B., Ott, G., Arnold, L. & Sprinzl, M. (1996). NMR evidence for helix geometry modifications by a G-U wobble base pair in the acceptor arm of E. coli $tRNA^{Ala}$. *FEBS Lett* **385**, 15-20.

Lynch, D. C. & Schimmel, P. R. (1974). Cooperative binding of magnesium to transfer ribonucleic acid studied by a fluorescent probe. *Biochemistry* **13**, 1841-1852.

Madhani, H. D. & Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* **28**, 1-26.

Marino, J. P., Gregorian, R. S. J., Csankovszki, G. & Crothers, D. M. (1995). Bent helix formation between RNA hairpins with complementary loops. *Science* **268**, 1448-1454.

McDowell, J. A. & Turner, D. H. (1996). Investigation of the structural basis for thermodynamic stabilities of tandem G-U mismatches: solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing. *Biochemistry* **35**, 14077-14089.

Michel, F., Ellington, A. D., Couture, S. & Szostak, J. W. (1990). Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. *Nature* **347**, 578-580.

Michel, F., Jacquier, A. & Dujon, B. (1982). Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* **64**, 837-881.

Michel, F. & Westhof, E. (1990). Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585-610.

Miranker, A. D., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. (1993). Detection of transient protein folding populations by MS. *Science* **262**, 896-900.

Mirmira, S. R. & Tinoco, I. J. (1996). NMR structure of a bacteriophage T4 RNA hairpin involved in translational repression. *Biochemistry* **35**, 7664-7674.

Moazad, D. & Noller, H. F. (1990). Binding of tRNA to the ribosomal A and P site protects two distinct sets of nucleotides in 16S rRNA. *J. Mol. Biol.* **211**, 135-145.

Moazad, D., Robertson, J. M. & Noller, H. F. (1988). Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S ribosomal RNA. *Nature* **334**, 362-364.

Moazed, D. & Noller, H. F. (1989). Interactions of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell* **57**, 585-597.

Moazed, D., Stern, S. & Noller, H. F. (1986). Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**, 399-416.

Mohr, G., Caprara, M. G., Guo, Q. & Lambowitz, A. M. (1994). A tyrosyl-tRNA synthetase can function similarly to an RNA structure in the *Tetrahymena* ribozyme. *Nature* **370**, 147-150.

Mohr, G., Zhang, A., Gianelos, J. A., Belfort, M. & Lambowitz, A. M. (1992). The *Neurospora* CYT-18 protein suppresses defects in the phage T4 *td* intron by stabilizing the catalytically active structure of the intron core. *Cell* **69**, 483-494.

Moore, P. B. (1993). Ribosomes and the RNA world. In *The RNA World* (Gesteland, R. F. & Atkins, J. F., eds.), pp. 119-135. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Murphy, F. L. & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry* **32**, 5291-5300.

Murphy, F. L. & Cech, T. R. (1994). GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236**, 49-63.

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Resketnikova, L., Clark, B. F. C. & Nyborg, J. (1995). Crystal structure of the ternary complex Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* **270**, 1464-1472.

Noller, H. F. (1991). Ribosomal RNA and translation. *Annu. Rev. Biochem.* **60**, 191-227.

Noller, H. F., Hoffarth, V. & Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416-1419.

Oliveberg, M., Tan, Y.-J. & Fersht, A. R. (1995). Negative activation enthalpies in the kinetics of protein folding. *Proceedings of the National Academy of Science* **92**, 8926-8929.

Orgel, L. E. (1968). Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381-393.

Pan, J., Thirumalai, D. & Woodson, S. A. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **293**, 7-13.

Pan, T. (1995). Higher order folding and domain analysis of the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry* **34**, 902-909.

Pan, T. & Sosnick, T. (1997). Intermediates and kinetic traps in the folding of a large ribozyme revealed by CD and UV spectroscopies and catalytic activity. *Nature Struct. Biol.* **14**(11), 931-938.

Peterson, E. R., Pan, T., Coleman, J. & Uhlenbeck, O. C. (1994). *In vitro* selection of small RNAs that bind to *Escherichia coli* phenylalanyl-tRNA synthetase. *J. Mol. Biol.* **242**, 186-192.

Pley, H., Flaherty, K. M. & McKay, D. B. (1994a). Three-dimensional structure of a hammerhead ribozyme. *Nature* **372**, 68-74.

Pörschke, D. & Eigen, M. (1971). Cooperative non-enzymatic base recognition III. Kinetics of the helix-coil transition of the oligoribouridylic-oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.* **62**, 361-381.

Privalov, P. L., Filimonov, V. V., Venkstern, T. B. & Bayev, A. A. (1975). A calorimetric investigation of tRNA$_1^{Val}$ melting. *J. Mol. Biol.* **97**, 279-288.

Puglisi, J., Wyatt, J. & Tinoco, I. J. (1990). Conformation of an RNA psuedoknot. *J. Mol Biol* **214**, 437-453.

Puglisi, J. D., Chen, L., Blanchard, S. & Frankel, A. D. (1995). Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **270**, 1200-1203.

Pyle, A. M. & Cech, T. R. (1991). Ribozyme recognition of RNA by tertiary interactions with specific ribose 2'-OH groups. *Nature* **350**, 628-631.

Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302-307.

Reich, C., Olson, G. J., Pace, B. & Pace, N. R. (1988). Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science* **239**, 178-181.

Robertus, J. D., Ladner, J. E., Finch, J. R., Thodes, D., Brown, R. S., Clark, B. F. C. & Klug, A. (1974). Structure of yeast phenylalanine tRNA at 3 A resolution. *Nature* **250**, 546-551.

Romer, R. & Hach, R. (1975). tRNA conformation and magnesium binding: A study of yeast phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. *Eur. J. Biochem.* **55**, 271-284.

Rothwarf, D. M. & Scheraga, H. A. (1996). Role of non-native aromatic and hydrophobic interactions in the folding of hen egg white lysozyme. *Biochemistry* **35**, 13797-13807.

Rould, M. A., Perona, J. J. & Steitz, T. A. (1991). Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. *Nature* **352**, 213-217.

Saldanha, R. J., Patel, S. S., Surendran, R., Lee, J. C. & Lambowitz, A. M. (1995). Involvement of *Neurospora* mitochondrial tyrosyl-tRNA synthetase in RNA splicing: A new method for purifying the protein and characterization of physical and enzymatic properties pertinent to splicing. *Biochemistry* **34**, 1275-1287.

Sassanfar, M. & Szostak, J. W. (1993). An RNA motif that binds ATP. *Nature* **364**, 550-553.

Schulze, H. & Nierhaus, K. H. (1982). Minimal set of ribosomal components for reconstitution of the peptidyltransferase activity. *EMBO J* **1**, 609-613.

Sclavi, B., Brenowitz, M. & Woodson, S. (1998). RNA folding at millisecond intervals by synchrotorn hydroxyl radical footprinting. *Science* **279**, 1940-1943.

Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L. & Klug, A. (1996). Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science* **274**, 2065-2069.

Scott, W. J., Finch, J. T. & Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81**, 991-1002.

Shen, L. X. & Tinoco, I. J. (1995). The structure of an RNA pseudoknot that causes efficient framshifting in mouse mammary tumor virus. *J Mol Biol* **247**, 963-978.

Shine, J. & Dalgarno, L. (1974). The 3' terminal sequence of Escherichia coli 16S ribosomal RNA gene sequences. *Nucleic Acids Res.* **19**, 2189-2191.

Simpson, L. & Thiemann, O. H. (1997). mRNA editing. In *Eukaryotic mRNA processing* (Krainer, A. R., ed.), pp. 335-369. IRL Press, Cold Spring Harbour.

Söll, D. (1993). Transfer RNA: an RNA for all seasons. In *The RNA World* (Gesteland, R. F. & Atkins, J. F., eds.), pp. 157-184. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994). The barriers in protein folding. *Nature Struct. Biol.* **1**, 149-156.

Steitz, J. A. (1992). Splicing takes a holiday. *Science* **257**, 888.

Stern, S., Weiser, B. & Noller, H. F. (1988). Model for the three-dimensional folding of 16S ribosomal RNA. *J. Mol. Biol.* **204**, 447-481.

Strobel, S. A. & Cech, T. R. (1993). Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the *Tetrahymena* ribozyme. *Biochemistry* **32**, 13593-13604.

Strobel, S. A. & Cech, T. R. (1996). Exocyclic amine of the conserved G•U pair at the cleavage site of the *Tetrahymena* ribozyme contributes to 5'-splice site selection and transition state stabilization. *Biochemistry* **35**, 1201-1211.

Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., Moncoeur, E. (1998). Complementary sets of noncononical base pairs mediate RNA helix packing in the group I intron active site. *Nat Struct Biol* **5**, 60-65.

Symons, R. H. (1992). Small catalytic RNAs. *Annu. Rev. Biochem.* **61**, 641-671.

Tan, Y., Oliveberge, M. , Fersht, A.R. Titration properties and thermodynamics of the transition state for folding: comparison of two-state and multi-state folding pathways. *J. Mol. Biol.* 264, 377-389.

Tinoco, I. J., Sauer, K. & Wang, J. C. (1995). *Physical Chemistry: Principles and Applications In Biological Sciences*. 3rd edit, Prentice Hall.

Treiber, D. K., Rook, M. S., Zarrinkar, P. & Williamson, J. R. (1998). Kinetic intermediates trapped by native interactions in RNA folding. *Science* **279**, 1943-1947.

Waring, R. B., Towner, P., Minter, S. J. & Davies, R. W. (1986). Splice-site selection by a self-splicing RNA of *Tetrahymena*. *Nature* **321**, 133-139.

Weeks, K. M. & Cech, T. R. (1995). Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5' splice site domain. *Cell* **82**, 221-230.

Weeks, K. M. & Cech, T. R. (1996). Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science* **271**, 345-348.

Westhof, E. & Altman, S. (1994). Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5133-5137.

Wildegger, G. & Keifhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *J. Mol. Biol.* **270**, 294-304.

Wilson, C. & Szostak, J. W. (1995). *In vitro* evolution of a self-alkylating ribozyme. *Nature* **374**, 777-782.

Woese, C. (1967). The evolution of the genetic code. In *The genetic code*, pp. 179-195. Harper and Row, New York.

Wolfson, J. M. & Kearns, D. R. (1975). Europium as a fluorescent probe of transfer RNA structure. *Biochemistry* **14**, 1436-1444.

Woodson, S. A. & Cech, T. R. (1991). Alternative secondary structures in the 5' exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry* **30**, 2042-2050.

Wu, M. & Turner, D. H. (1996). Solution structure of $(rGCGGACGC)_2$ by two-dimensional NMR and the iterative relaxation matrix approach. *Biochemistry* **35**, 9677-9689.

Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science* **272**, 1343-1347.

Yu, Y.-T., Maroney, P. A., Darzynkiewicz, E. & Nilsen, T. W. (1995). U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* **1**, 46-54.

Zarrinkar, P. P., Wang, J. & Williamson, J. R. (1996). Slow folding kinetics of RNase P RNA. *RNA* **2**, 564-573.

Zarrinkar, P. P. & Williamson, J. R. (1994). Kinetic intermediates in RNA folding. *Science* **265**, 918-924.

Zarrinkar, P. P. & Williamson, J. R. (1996). The kinetic folding pathway of the *Tetrahymena* ribozyme reveals possible similarities between RNA and protein folding. *Nature Struct. Biol.* **3**(5), 432-438.

Zaug, A. J., Been, M. D. & Cech, T. R. (1986). The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease. *Nature* **324**, 429-433.

Zaug, A. J. & Cech, T. R. (1982). The intervening sequence excised from the ribosomal RNA precursor of *Tetrahymena* contains a 5'-terminal guanosine residue not encoded by the DNA. *Nucleic Acids Res.* **10**, 2823-2838.

Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988). Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry* **27**, 8924-8931.

Zhang, A., Derbyshire, V., Galloway Salvo, J. L. & Belfort, M. (1995a). *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA* **1**, 783-793.

Zhang, F., Ransay, E. S. & Woodson, S. A. (1995b). *In vivo* facilitation of *Tetrahymena* group I intron splicing in Escherichia coli pre-ribosomal RNA. *RNA* **1**, 284-292.

Biographical Note

Martha S. Rook

Education

Sept. 1992-June 1998	Ph.D. in Biological Chemistry: Massachusetts Institute of Technology Department of Chemistry Advisor: Prof. James R. Williamson
August 1988-May 1992	Bachelor of Science in Chemistry Cum Laude Texas A&M University

Awards and Honors

1993-1994	Texaco Graduate Fellowship
1991-1992	GE Foundation Faculty for the Future Scholarship Academic Excellence Award Undergraduate Chemistry Achievement Award Barnes and Noble Scholarship
1989	Undergraduate Chemistry Achievement Award Lechner Merit Award
1988-1990	Arthur E. Martell IUCCP Scholarship

Experience

1992-1998	Research Assistant, Massachusetts Institute of Technology Research Advisor: Associate Professor James R. Williamson
1992-1993	Teaching Assistant, Massachusetts Institute of Technology Introductory Chemistry Introductory Laboratory Chemistry Research focused on kinetic folding of the *Tetrahymena* ribozyme.
1990-1992	Undergraduate Research Assistant, Texas A&M University Research Advisor: Associate Professor Jeffrey W. Kelly Research focused on the study of the subunit assembly of the transthyretin protein .

1988-1992 Summer Research	Undergraduate Research Assistant United States Medical Research Institute of Infectious Diseases Fort Detrick, Frederick, MD Research Advisor: Lt. Col. John Lowe Cloning and sequencing of genes from *Bacillus anthracis*.

Publications

Treiber, D.K., Rook, M.S., Zarrinkar, P.A., Williamson, J.R. Kinetic intermediates trapped by native interactions in RNA folding. *Science,* **279**, 1943-1946.

Rook, M.S., Treiber, D.K., Williamson, J.R. "Fast Folding Mutants of the *Tetrahymena* Ribozyme Reveal a Rugged Folding Energy Landscape", *submitted.*

Presentations

Rook, M.S., Williamson, J.R. "Fluorescence Studies of the Kinetic Folding Pathway of the *Tetrahymena* Group I Intron" Biophysical Society 40th Annual Meeting Poster Session, Baltimore, MD, 1996.